

Modelo de Ecuaciones Estructurales para la evaluación de la calidad de una vacuna bovina

*Tesis presentada para optar al título de Magister de la Universidad de Buenos Aires,
Área Biometría*

María Virginia López

Ingeniera Agrónoma - Universidad de Buenos Aires- 1981

Lugar de trabajo: Facultad de Agronomía, Universidad de Buenos Aires



Escuela para Graduados Ing. Agr. Alberto Soriano
Facultad de Agronomía – Universidad de Buenos Aires

COMITÉ CONSEJERO

Director de tesis

Marta Beatriz Quaglino

Estadística (Universidad Nacional de Rosario)

Doctora en Matemática, especialidad Estadística (Universidad Politécnica de Valencia)

Consejero de Estudios

Gladys Viviana Parreño

Bioquímica (Universidad de Morón)

Doctora , área Microbiología e Inmunología Veterinaria (Universidad de Buenos Aires)

Consejero de Estudios

Norberto José Bartoloni

Ingeniero Agrónomo (Universidad de Buenos Aires)

Magister Scientiae, área Biometría (Universidad de Buenos Aires)

JURADO DE TESIS

Director de tesis

Marta Beatriz Quaglino

Estadística (Universidad Nacional de Rosario)

Doctora en Matemática, especialidad Estadística (Universidad Politécnica de Valencia)

JURADO

Antonio Humberto Closas

Licenciado en Matemáticas (Universidad Pública de Navarra)

Doctor (Universidad Pública de Navarra)

JURADO

Andrés Pérez

Médico Veterinario (Universidad de Buenos Aires)

Doctor (Universidad de Buenos Aires)

Fecha de defensa de la tesis: 5 de Diciembre de 2012

Agradecimientos

En primer lugar quiero agradecer a Dios, que es la fuente de vida, por haberme dado la salud y capacidad para llegar hasta aquí. A Él sea toda la honra y el honor.

A todas las personas e instituciones que colaboraron en mi formación a lo largo de todos estos años.

A mi esposo por su paciencia y aliento en cada etapa de esta tesis. Gracias, Jorge, por permitirme ser parte de este “equipo de dos” que pudo llegar a este logro.

A la Dra. Viviana Parreño por su entusiasmo contagioso y constante estímulo a lo largo de este trabajo. Si no fuera por vos, Vivi, nunca hubiera podido incursionar en el mundo de la virología animal y conocer el modelo cobayo.

A la Estadística Laura Marangunich, por haber compartido tan generosamente su experiencia estadística en el área de la salud animal. Gracias, Laura, por brindarme sin mezquindades todo tu saber.

A la Dra. Marta Quaglino, que aceptó el desafío de dirigir esta tesis y que ha hecho tan grato todo el tiempo de su elaboración. Muchas gracias, Marta, por todas las cosas que has tenido que dejar por acompañarme.

Al Ing. Agr. MSc. Norberto Bartoloni, por su constante apoyo y por sus consejos que han sido tan útiles. Gracias, Norberto, por tu amistad.

A la Dra. María del Carmen Fabrizio por su compañerismo en los momentos difíciles y por su permanente estímulo para terminar esta tesis. Muchas gracias, Mari, por compartir tantas alegrías y logros en este transitar por los senderos académicos.

A la Lic. MSc. Nora Abbiatti por su interés en que concluyera con éxito la maestría y por haber sido una de mis mejores docentes en el postgrado.

A la Dra. Alejandra Romera por haberme ayudado con los datos de protección, indispensable para este trabajo.

Finalmente, pero no menos importante, a toda mi familia, por su amor incondicional y paciencia.

Declaro que el material incluido en esta tesis es, a mi mejor saber y entender, original producto de mi propio trabajo (salvo en la medida en que se identifique explícitamente las contribuciones de otros), y que este material no lo he presentado, en forma parcial o total, como una tesis en ésta u otra institución.

ÍNDICE GENERAL

INTRODUCCIÓN	1
CAPÍTULO 1: EL PROBLEMA BIOLÓGICO	7
1.1. Rinotraqueítis Infecciosa Bovina	7
1.2. Vacunas	8
1.3. Calidad de las vacunas	12
CAPÍTULO 2: FUNDAMENTOS DE LOS MODELOS DE ECUACIONES ESTRUCTURALES	15
2.1. Ecuaciones estructurales	15
2.2. Relación entre variables latentes y observadas	16
2.3. Componentes del sistema y notación	17
2.3.1. Modelo de variable latente	17
2.3.2. Modelo de medida	19
2.4 El modelo general	21
2.5 Análisis de Senderos	22
2.5.1 Diagrama de senderos	22
2.6 Identificación del modelo	24
2.6.1 Regla	24
2.6.2 Otras técnicas	24
2.7 Estimación de los parámetros del modelo	26
2.7.1 Métodos de estimación	26
2.7.2 Soluciones numéricas para minimizar las funciones de ajuste	28

2.7.2.1 Valores iniciales	29
2.7.2.2 Pasos en la secuencia	33
2.7.2.3 Criterio de finalización	35
2.8. Errores estándar aproximados	35
2.9 Bondad de Ajuste del Modelo	35
2.9.1 Medidas de ajuste global	36
2.9.1.1 Residuales	36
2.9.1.2 Prueba de χ^2	38
2.9.1.3 Índices de ajuste incremental	40
2.9.1.4. Otras medidas de ajuste	42
2.9.2 Medidas de Ajuste de los componentes del modelo	44
2.9.2.1 Estimadores de los parámetros	44
2.9.2.2 Prueba de hipótesis para los coeficientes	45
2.9.2.3 Matriz de correlación asintótica de los estimadores de los parámetros	46
2.9.2.4 R^2 de las variables observadas	46
2.9.2.5. Fiabilidad compuesta	47
2.9.2.6 Variancia extraída	47
2.9.3 Comparación de modelos	47
2.9.3.1 Prueba del cociente de verosimilitud (LR)	48
2.9.3.2 Prueba del Multiplicador de Lagrange (LM)	48
2.9.3.3. Prueba de Wald (W)	49
2.9.3.4. Medida de ajuste de la componente estructural del modelo	50

CAPÍTULO 3: ESPECIFICACION DEL MODELO ESTRUCTURAL PARA LA VALIDACIÓN DE LA EVALUACIÓN DE VACUNAS CONTRA BoHV-1	51
3.1. Variables del modelo	51
3.2. El modelo de ecuaciones estructurales para la validación de un método de evaluación de vacunas	53
3.3. Experimentos y medición de variables	58
3.4. Procedimiento de simulación	63
3.5. Estrategia de análisis	63
3.6. Introducción de datos	64
CAPÍTULO 4: RESULTADOS	65
4.1. Modelo de medida	65
4.1.1. Modelo de medida sin modificaciones	65
4.1.2. Modificaciones del modelo de medida	69
4.1.3 Modelo de medida final	72
4.1.3.1 Medidas de ajuste global	72
4.2 Medidas de Ajuste de los componentes del modelo	75
4.2.1 Estimadores de los parámetros	75
4.2.2 Otras medidas de evaluación de los componentes del modelo de medida	78
4.3 Modelo estructural	79
4.3.1 Ajuste global del modelo estructural	80
4.3.2 Ajuste de los componentes del modelo	81
4.4 Ajuste de la componente estructural del modelo	83
4.5 Resultado de simulaciones de 15 muestras	84

	xii
CAPÍTULO 5: DISCUSIÓN	87
CAPÍTULO 6: CONCLUSIONES	93
APÉNDICE	95
A.1. Programa para simular los valores serológicos de cobayos a partir de los valores bovinos	95
A2 Programa para generar la muestra Normal multivariada	95
A3. Programa para el análisis del modelo de ecuaciones estructurales para la validación de un método alternativo de evaluación de vacunas	98
BIBLIOGRAFÍA	101

ÍNDICE DE CUADROS

Cuadro 2.1: Notación para modelos de variables latentes	19
Cuadro 2.2: Notación para modelos de medida	20
Cuadro 2.3: Símbolos principales empleados en el Análisis de senderos	23
Cuadro 3.1: Medidas descriptivas de las variables manifiestas en la base de datos original	60
Cuadro 3.2: Resultados del análisis de regresión de las variables serológicas de los cobayos en función de bovinos del Campo 3 (animales no respondedores) y del resto	60
Cuadro 3.3.: Resultados del análisis de regresión con los valores serológicos de cobayos simulados	61
Cuadro 3.4.: Resultados del análisis de regresión con los valores serológicos de cobayos simulados	62
Cuadro 4.1: Comparación de modelos de medidas de evaluación global	72
Cuadro 4.2: Medidas de ajuste global para cada componente del modelo	74
Cuadro 4.3: Medidas de ajuste global para los modelos con las variables latentes de a pares	75
Cuadro 4.4: Propiedades del modelo de medida	78
Cuadro 4.5: Resumen del ajuste de modelos	83
Cuadro 4.6: Índices de ajuste global del modelo estructural resultados de 15 muestras extraídas aleatoriamente de la población simulada	84
Cuadro 4.7: Coeficientes estandarizados y fiabilidades de las variables según el modelo estructural resultado de 15 muestras extraídas de una población simulada	85

ÍNDICE DE FIGURAS

Figura 3.1: Diagrama de senderos de las relaciones entre variables en la evaluación de vacunas contra el virus BoHV-1	55
Figura 4.1: Índices de ajuste global del modelo de medida inicial	65
Figura 4.2: Resultado de la salida del programa SAS con los 10 valores de residuales estandarizados absolutos más altos	66
Figura 4.3: Coeficientes de las variables observadas y estadístico t	67
Figura 4.4: Fiabilidad de las variables indicadoras en el modelo inicial	68
Figura 4.5: Prueba de Wald para el modelo inicial	68
Figura 4.6: Multiplicadores de Lagrange en orden descendente de importancia	69
Figura 4.7: Índices de ajuste global del modelo de medida con correlación de errores	69
Figura 4.8: Fiabilidad de las variables indicadoras en el modelo con correlación de errores	70
Figura 4.9: Multiplicadores de Lagrange en orden descendente de importancia para el modelo con correlación de errores	71
Figura 4.10: Índices de ajuste global del modelo de medida final	73
Figura 4.11: Salidas de SAS de las ecuaciones con las variables manifiestas y sus estimaciones en el modelo de medida	76
Figura 4.12: Salida del software con el resumen de las medidas de ajuste global del modelo estructural	80
Figura 4.13: Salida del <i>software</i> con los 10 residuales normalizados extremos	81
Figura 4.14: Salida del <i>software</i> con las ecuaciones de las variables manifiestas y latentes endógenas en el modelo estructural	81
Figura 4.15: Salida del <i>software</i> con los valores de R^2 de las variables endógenas y los coeficientes de correlación de las variables exógenas	82
Figura 4.16: Diagrama de senderos final y coeficientes estandarizados estimados	84

RESUMEN

Los modelos de ecuaciones estructurales fueron introducidos en el contexto de la Genética y la Econometría y han sido extensamente empleados en las Ciencias Sociales. Sin embargo han sido poco empleados en el área de Salud Animal. En este trabajo se propone la construcción de un modelo de ecuaciones estructurales para la validación de un sistema de evaluación de la calidad de una vacuna contra la Rinotraqueítis infecciosa bovina. La dificultad para encontrar rodeos de bovinos seronegativos libres del virus BoHV-1 plantea la necesidad de desarrollar pruebas estandarizadas y armonizadas en animales de laboratorio que permitan definir la consistencia de lote a lote de la potencia de la vacuna, de un modo rápido y menos costoso para asegurar la presencia de productos confiables en el mercado. En este sentido, es de interés verificar la concordancia de la respuesta serológica en el animal de laboratorio con la protección a la enfermedad y a la infección en bovinos. Este sistema involucra variables latentes y relaciones complejas entre variables. El modelo de ecuaciones estructurales demostró ser una técnica útil para la validación de un modelo que emplea mediciones serológicas en animales de laboratorio como alternativa a las pruebas clásicas de desafío en bovinos para la evaluación de la efectividad de las vacunas empleadas en el control y prevención de la enfermedad causada por un herpesvirus.

En este trabajo se pudo comprobar que las variables que representaban los conceptos de interés - *Protección a la infección*, *Protección a la enfermedad* y *Calidad determinada por el modelo animal en el laboratorio* (medición alternativa de la efectividad de la vacuna) explicaron adecuadamente la variabilidad de las variables observables. Se destacaron las variables *Área bajo las curvas de excreción* y *Área bajo la curva de severidad* para representar las variables *Protección a la infección* y *Protección a la enfermedad*, respectivamente, y las *variables serológicas medidas por ELISA* en bovinos y cobayos representando la *Calidad por laboratorio*. Por otra parte, el modelo permitió la incorporación de correlación de errores de las variables serológicas, que hubiese sido imposible con otras técnicas multivariadas. Asimismo, la técnica de análisis permitió explorar por partes las relaciones entre variables para detectar la fuente de falta de ajuste. Se pudo comprobar el buen ajuste de las relaciones de las variables latentes *Protección a la infección* y *Calidad por laboratorio*, obteniéndose una alta correlación entre ambos conceptos. Teniendo en cuenta la alta fiabilidad de las variables serológicas medidas por ELISA tanto en bovinos como en cobayos, se puede inferir que la medición de estas variables de laboratorio puede ser útil en la predicción de la protección de la vacuna. El efecto de la *Protección a la infección* sobre la *Protección a la enfermedad* fue significativo pero débil. La parte estructural del modelo presentó un ajuste adecuado, validando así la teoría inicial de interrelaciones entre las variables latentes. Se plantea la posibilidad de extender los modelos de ecuaciones estructurales para la evaluación de vacunas que involucran otros virus que afectan el ganado bovino.

Palabras clave: modelo de ecuaciones estructurales, modelo de variable latente, modelo de medida, potencia de la vacuna, Rinotraqueítis infecciosa.

ABSTRACT

Structural equation models were introduced in the context of Genetics and Econometrics and have been widely used in Social Sciences. Yet they have been little used in the area of animal health. This thesis proposes the construction of a structural equation model for the validation of a system for quality assessment of a vaccine against Infectious Bovine Rhinotracheitis. The difficulty in finding cattle herds free of the herpesvirus BoHV-1 raises the need to develop harmonized and standardized tests in laboratory animals to define the consistency from batch to batch of vaccine potency, in a fast and less expensive way to ensure the presence of reliable products on the market. In this regard, it is of interest to verify the consistency of the serological response of laboratory animals with the protection from disease and infection in cattle. This system involves latent variables and complex relationships between variables. The structural equation model proved to be a useful technique to validate a model that used serological measurements in laboratory animals as an alternative to the classical challenge tests in cattle to evaluate the effectiveness of the vaccines used in the prevention and control of a disease caused by a herpes virus. It was found that variables representing the concepts of interest that is *Protection to Infection*, *Protection to Disease* and *Laboratory Quality* (alternative measure of effectiveness of the vaccine) adequately explain the variability of the observed variables. The observed variables *Area Under Excretion Curve* and *Area Under Severity Curve* were the most important to represent the latent variables *Protection to Infection* and *Protection to Disease*. In the same manner, the observed serological variables measured by ELISA for both animal species were the most important to represent *Laboratory Quality*. Moreover, the model allowed the incorporation of correlated measurement errors which would have been impossible with other multivariate techniques. Also, this methodology allowed exploring the relationship between variables by components to detect the source of lack of fit. It was found that the relationship between the latent variables *Protection to Infection* and *Laboratory Quality* presented a good fit, showing a high correlation between these concepts. Considering the high reliability of the serological variables measured by ELISA for both animal species, it can be inferred that the measurement of these laboratory variables may be useful in predicting vaccine protection. The effect of the latent variable *Protection to Infection* on *Protection to Disease* was weak but significant. The structural component of the model showed a good fit, thus validating the initial theory of the relationships between latent variables. The possibility of extending the structural equation modeling for vaccine evaluation involving other viruses that affect cattle is suggested.

Key words: structural equation model, latent variable model, measurement model, vaccine potency, Infectious Bovine Rhinotracheitis.

INTRODUCCIÓN

El herpesvirus 1 del bovino (BoHV-1), agente causal de la Rinotraqueítis Infecciosa Bovina (IBR), constituye una causa importante de pérdidas económicas en la actividad ganadera en todo el mundo (Gibbs y Rweyemamu, 1997). Este virus afecta al ganado bovino causando síndromes respiratorios, reproductivos y queratoconjuntivitis. Dentro de los signos clínicos, la infección por BoHV-1 provoca inapetencia, rinitis, tos, vulvovaginitis, balanopostitis, caída de la producción de leche e inmunosupresión favoreciendo infecciones por parte del patógeno oportunista. La infección se transmite en forma vertical de las hembras preñadas al feto causando abortos y mortalidad neonatal. Al igual que otros herpesvirus, BoHV-1 puede causar infecciones que persisten por largo tiempo en forma latente en los ganglios nerviosos sensoriales y que ante una situación de estrés pueden sufrir una reactivación y los animales pueden excretar virus asintóticamente.

Luego de la implementación de programas estrictos de control, la enfermedad ha sido erradicada de los países europeos nórdicos (Noruega, Finlandia y Suecia), Austria, Dinamarca, y parte de Italia. Actualmente, otros países europeos están bajo programas de erradicación obligatorios o voluntarios, involucrando todos ellos la aplicación de vacunas inactivadas o vacunas “marcadoras” basadas en la delección de las glicoproteínas virales BoHV-1 gE o gD. En el resto del mundo, todavía están disponibles las clásicas vacunas con BoHV-1 muerto o atenuado (Ackermann y Engels, 2006).

En Argentina, la infección por BoHV-1 es endémica y sólo se permite el uso de vacunas con virus inactivados para prevenir la enfermedad (SENASA, Resol. 216/2006; Resol. 351/0607). Varias vacunas combinadas que contienen BoHV-1 junto con otros virus y bacterias patógenas están disponibles en el mercado local. Estas vacunas multivalentes han sido diseñadas para controlar un problema sanitario de etiología compleja. Sin embargo, la eficacia y potencia¹ contra cada uno de los antígenos contenidos en algunas de estas formulaciones combinadas no están claras y requieren ser evaluadas correctamente.

Para determinar el nivel de anticuerpos alcanzado por los animales infectados o vacunados con BoHV-1 (BoHV-1Ab), varios ELISAs² fueron desarrollados y probados como más sensibles y específicos que la prueba de neutralización viral (SN)³ (Kramps *et al.*, 2004; OIE, 2008). Sin embargo, las agencias regulatorias no han adoptado todavía ninguno de los ELISA desarrollados para la aprobación de vacunas y SN sigue siendo considerada la técnica “Gold Standard” para la validación de la prueba de potencia de vacunas (CFR 1985; World Organization for Animal Health, 2008).

¹ Potencia de una vacuna es su capacidad para producir una respuesta inmunitaria determinada.

² La técnica **ELISA** (*Enzyme-Linked ImmunoSorbent Assay*) se basa en la detección de un antígeno inmovilizado sobre una fase sólida mediante anticuerpos que directa o indirectamente producen una reacción cuyo producto, por ejemplo un colorante, puede ser medido por espectrofotometría.

³ La neutralización viral se define como la pérdida de la infectividad de un virus, por la reacción del mismo con un anticuerpo específico.

Específicamente, para la infección por BoHV-1, los organismos regulatorios internacionales (CFR, 1985; EMEA, 1998, 1999; OIE, 2008) recomiendan evaluar la inmunogenicidad con una prueba de potencia conducida en terneros seronegativos. Cada serie o subserie del lote de vacunas con virus BoHV-1 inactivado debería ser probado en terneros susceptibles a la infección por IBR (negativos para anticuerpos neutralizantes a BoHV-1). Luego de la vacunación los animales controles deberían permanecer seronegativos y al menos 4 de 5 terneros vacunados deberían desarrollar títulos neutralizantes Ab de 1:8 o mayores. Las vacunas que no alcanzan este requerimiento de inmunogenicidad deberían someterse a una prueba de desafío por vía intranasal. En esta prueba, dos de tres terneros controles deberían desarrollar signos de hipertermia y signos clínicos de IBR, mientras que sólo uno de 5 terneros vacunados podría desarrollar signos clínicos; de otro modo el lote de vacuna es considerado de pobre inmunogenicidad [CFR 1985 9.113.216. 55 FR 35562, Agosto 31, 1990, enmendado en 56 FR 66786, Diciembre 26, 1991]. Posteriormente al desafío experimental, la OIE también requiere que la vacuna reduzca la descarga nasal de virus al menos 100 veces y el período de excreción en tres días comparando con los terneros controles (World Organization for Animal Health, 2008). Así, para comercializarla, una vacuna a BoHV-1 debe prevenir el desarrollo de síntomas clínicos severos y reducir marcadamente la excreción de virus.

Los ensayos con bovinos son costosos e insumen gran cantidad de tiempo para ser empleados como pruebas de rutina para liberar al mercado un lote de vacuna, especialmente en países como Argentina donde BoHV-1, así como otras infecciones virales tales como diarrea viral bovina, parainfluenza bovina tipo 3 y rotavirus, son altamente endémicas (Campero *et al.*, 2003; Moore *et al.*, 2003). La dificultad en encontrar rodeos de bovinos seronegativos libres de BoHV-1 para usarlos en la prueba de potencia de vacunas, plantea la necesidad de desarrollar pruebas estandarizadas y armonizadas en animales de laboratorio que permitan definir la consistencia de lote a lote de la potencia de la vacuna, de un modo más rápido y menos costoso para asegurar la presencia de productos confiables en el mercado. La Farmacopea Europea ha respaldado el uso de métodos de prueba alternativos, siempre que pueda establecerse una correlación entre ellos y la prueba del desafío. El principio III de los criterios regulatorios para la revisión de protocolos definidos en PHS Policy y las regulaciones de USDA (Institutional Animal Care and use committee Guidebook, 2002) dice: “Los animales seleccionados por un procedimiento deben ser de una especie y calidad apropiadas y requerir un mínimo de animales para obtener resultados válidos. Métodos tales como modelos matemáticos, simulación por computación, y sistemas biológicos *in vitro* deben considerarse”.

Si bien algunos productores de vacunas han empleado animales de laboratorio para probar sus vacunas (Merial ®) por medio de protocolos internos, sólo se tiene conocimiento de un trabajo sobre la evaluación de un modelo en cobayo para estudiar la inmunogenicidad de vacunas en Herpesvirus del bovino (Alonzo *et al.*, 2008). Por otra parte, el Instituto Nacional de Tecnología Agropecuaria (INTA) en convenio con el laboratorio Biogénesis Bagó ha trabajado en este sentido, validando un método alternativo para la prueba de potencia de vacunas para BoHV-1 en cobayos (Parreño *et al.*, 2010). La prueba fue especialmente diseñada para evaluar la inmunogenicidad a BoHV-1 y otros antígenos virales dentro del grupo de vacunas combinadas inactivadas (Taffs, 2001). Fue desarrollada en los productos terminados en un modelo *in vivo*, con el objetivo de detectar no sólo la variabilidad en cantidad y calidad del antígeno sino también las variables que afectan la inmunogenicidad, tales como la potencia del

adyuvante, la estabilidad, y la variabilidad durante el proceso de formulación (McVey *et al.*, 2003).

En el trabajo de Parreño *et al.* (2010) se probaron vacunas con distinto contenido de antígenos en bovinos y cobayos midiéndose variables relacionadas con la respuesta de anticuerpos en bovinos y cobayos, y relacionadas con la protección a la infección y a la enfermedad en bovinos. Los análisis estadísticos para la validación del modelo cobayo para evaluar la calidad de las vacunas destinadas a bovinos fueron univariados en este estudio. Sin embargo, dada la complejidad de las relaciones entre las variables involucradas en el modelo, un método multivariado resulta apropiado. En este sentido, el modelo de ecuaciones estructurales (SEM) se presenta atractivo básicamente por dos razones: (1) porque proporciona un método directo de tratar con múltiples relaciones simultáneamente a la vez que se da eficacia estadística, y (2) por su capacidad para evaluar las relaciones exhaustivamente y proporcionar una transición desde el **análisis exploratorio** al **confirmatorio** (Hair *et al.*, 2007). Las técnicas SEM se distinguen por dos características. Por un lado, a diferencia de otras técnicas multivariadas, SEM puede emplear relaciones distintas para cada conjunto de variables dependientes. En términos sencillos, SEM estima una serie de ecuaciones de regresión múltiple distintas pero interrelacionadas mediante la especificación del modelo estructural utilizado. Por otro lado, SEM también tiene la habilidad de incorporar variables latentes al análisis. Una variable latente es un concepto supuesto y no observado que sólo puede ser aproximado mediante variables medibles u observables. La incorporación de variables latentes tiene justificaciones tanto prácticas como teóricas, al mejorar la estimación estadística, mejorar la representación de los conceptos teóricos y tener en cuenta el error de medida. Desde el punto de vista práctico y teórico se sabe que no se puede medir perfectamente un concepto y que siempre hay un grado de **error de medida** (que no se tiene en cuenta en otras técnicas multivariadas).

En el desarrollo de un modelo para la evaluación de la potencia de una vacuna para BoVH1 se busca verificar la concordancia de la respuesta de anticuerpos en cobayos con la protección a la enfermedad y a la infección en bovinos, y, a su vez, con la respuesta de anticuerpos en bovinos. Este sistema involucra variables latentes ya que las variables *protección a la enfermedad*, *protección a la infección* y *calidad de vacuna* asignada por el laboratorio como evaluación alternativa de la protección, no son directamente observables. La situación ideal es lograr protección a la infección y a la enfermedad evaluando la calidad de la vacuna a partir del nivel de anticuerpos alcanzado por los bovinos y, mejor aún, por los cobayos. Un animal protegido es aquel que resulta refractario a la infección. En estos casos se dice que la inmunidad conferida por la vacuna es esterilizante ya que no permite el ingreso del virus al organismo, mientras que un animal protegido contra la enfermedad es aquel animal vacunado, que al estar expuesto al virus se infecta, el agente se multiplica en su organismo y se establece la infección latente y puede ser transmisor del virus; sin embargo, no presenta síntomas de la enfermedad. El primer tipo de protección, si bien constituye el ideal de vacuna, es difícil de alcanzar. Actualmente no existen vacunas para controlar IBR que sean esterilizantes. Por esta razón, no se puede considerar la protección como un concepto dicotómico en el caso de IBR sino que resulta más adecuado referirse a grados o niveles de protección.

El modelo de ecuaciones estructurales (SEM) abarca una familia entera de modelos conocidos con muchos nombres, entre ellos *análisis de la estructura de la covarianza*, *análisis de variable latente*, *análisis de factor confirmatorio* y a menudo simplemente *análisis LISREL* (el nombre de uno de los programas de *software* más populares). Producto de una evolución de la modelización multiecuacional, los modelos

de ecuaciones estructurales tienen una larga historia desarrollada principalmente en la econometría y fusionada con los principios de medición de la psicología y la sociología. Los métodos fueron introducidos en el contexto de la Genética y en Econometría y han sido extensamente empleados en las ciencias sociales. El genetista Sewall Wright, quien desarrolló el denominado “*path analysis*” en 1921 (Crow, 1988), y los economistas Trygve Haavelmo (1944) y Herber Simon (1952, 1954), fueron pioneros en el tema de las relaciones causales y las ecuaciones estructurales. Sin embargo, recién con el advenimiento de LISREL (Joreskog y Sorbom, 1989) fue posible el análisis de modelos causales con variables latentes en forma más amplia en diversas áreas del conocimiento, incluyendo la educación, el marketing, la psicología, la sociología, la gestión, la contrastación y medida, la salud, la demografía, el comportamiento organizacional. Austin (2007) menciona, entre otros trabajos de Ecología en los que SEM es aplicado, trabajos en Ecología Genética y Evolución (Mitchell, 1992, 1994), Ecofisiología Comparativa (Shipley y Lechowicz, 2000), interacciones tróficas (Marquez *et al.*, 2004), conservación de especies raras (Iriondo *et al.*, 2003). McCune y Grace (2002) proveen una detallada introducción de su uso en Ecología. No obstante, más recientemente, Armsworth *et al.* (2009) discuten el diferente enfoque de los economistas y ecólogos en el modo de analizar las relaciones entre variables, por lo que considera que la completa aplicación de la modelación por ecuaciones estructurales sigue siendo poco común en Ecología en comparación con Economía. En la literatura epidemiológica veterinaria, en el área de animales grandes, se ha propuesto el uso de *path analysis* para modelar desórdenes postparto de vacas lecheras junto con variables relacionadas con el ternero (Correa *et al.*, 1993). Más recientemente, Gianola y Sorensen (2004) propusieron un modelo de relaciones simultáneas entre fenotipos consideradas en el contexto de un modelo Gaussiano multi-carácter que estimuló la aplicación de SEM en el área del Mejoramiento Animal y Genética (de Los Campos *et al.* 2006a y 2006b; Varona *et al.* 2007; Wu *et al.*, 2007 y 2008; López de Marturana *et al.* 2007 y 2010; Heringstad *et al.*, 2008). No se tiene conocimiento del empleo del modelo de ecuaciones estructurales para la validación del método de evaluación de vacunas por una especie animal de laboratorio. No obstante, la utilización de animales de laboratorio, que permite obtener resultados extrapolables a menor costo que los animales destino, ha sido estudiada (Baba *et al.*, 1987; Goddard *et al.* 1986; Schleiss, 2008; Silva *et al.*, 2007; Smeets y Leonhardt, 1994; Yao *et al.*, 2008). La medición de los anticuerpos neutralizantes ha sido utilizada como método económico indirecto de evaluar la eficacia de las vacunas, basado en estudios que comparan el título de anticuerpos neutralizantes con pruebas de desafío para varios virus de interés veterinario (Meeusen *et al.*, 2007) y específicamente para BoHV (Patel, 2005). Con base en estos antecedentes, Alonzo *et al.* (2008) trabajaron con bovinos y cobayos a los que se les inyectó una vacuna comercial. La prueba de seroneutralización fue la empleada para estudiar inmunogenicidad del componente BoHV-1 de la vacuna en ambas especies. La cinética del nivel de anticuerpos en los animales fue analizada mediante un modelo de regresión, estimándose gráficamente los días en que el título fue máximo y en el que descendió por debajo de 8, que es el título considerado como protector por el Código Federal de Regulación de USDA (CFR, 2008). Los autores compararon las curvas de regresión de las dos especies animales.

Parreño *et al.* (2010), basaron la validación del modelo cobayo en las recomendaciones establecidas en la norma CVMP/IWP/038/97 de la **Agencia Europea de Evaluación de Productos Médicos, Unidad de Medicina Veterinaria (EMA)**. Como parte de su desarrollo se estableció la cinética de la respuesta inmune en el cobayo y se ajustó el volumen, número e intervalo entre las dosis para asemejar la

cinética de respuesta en el bovino. La validación de la prueba incluyó un ensayo dosis - respuesta para cada antígeno viral, en ambas especies el que, mediante un análisis de regresión, permitió definir la relación entre la dosis de antígeno (Ag) administrada y la respuesta en anticuerpos (Ac) evaluada a través de ELISA y SN. La validación estadística del modelo se basó en establecer la concordancia entre el modelo cobayo y la especie bovina, mediante métodos estadísticos paramétricos que incluyeron ensayos dosis respuesta, análisis de regresión y análisis de concordancia (índice Kappa). Por otra parte, se condujo un experimento de desafío en el que bovinos seronegativos fueron infectados por vía intranasal con el virus y fueron vacunados con vacunas clasificadas por el modelo cobayo y bovino como *no satisfactorias*, *satisfactorias* y *muy satisfactorias*, evaluándose síntomas clínicos y excreción de virus.

En el mercado nacional se encuentra disponible un gran número de vacunas destinadas a la prevención y control de enfermedades virales del complejo respiratorio y reproductivo del bovino. Estas formulaciones incluyen los siguientes agentes virales: Virus de la Rinotraqueítis Infecciosa Bovina (IBR), Virus Parainfluenza 3 (PI-3), Virus de la Diarrea Viral Bovina (BVDV) y virus respiratorio sincitial bovino (BRSV).

Particularmente en Argentina, para la comercialización de estos productos, el SENASA se basa en las pruebas de potencia ó eficacia realizadas en bovinos por cada laboratorio productor. Una vez aprobada la vacuna, cada serie es liberada luego de controles de esterilidad e inocuidad realizados por el SENASA, pero la calidad inmunogénica no es controlada: sólo se basa en las pruebas de potencia o inmunogenicidad en bovinos o en animales de laboratorio, que realizan los laboratorios productores utilizando protocolos internos propios (a través de declaraciones juradas) que no están armonizados. **En este sentido el SENASA ha manifestado la necesidad de incorporar una prueba de potencia en sus controles de rutina para estas vacunas combinadas.**

La obtención de un modelo estadístico que confirme una técnica de control de la calidad de vacuna, que ha sido analizada por otras técnicas cuantitativas, modelando las complejas relaciones entre las variables involucradas en el sistema, permitirá transferir una herramienta económica y de fácil implementación a las autoridades sanitarias como modelo de control de las vacunas polivalentes actualmente presentes en el mercado.

El **objetivo general** del presente trabajo es construir un modelo de ecuaciones estructurales para la validación de un sistema de evaluación de calidad de vacunas para una enfermedad bovina endémica en nuestro país.

Los **objetivos secundarios** son la valoración del sistema de ecuaciones estructurales en la evaluación de la fiabilidad de un nuevo método de medición y la comparación de dos métodos midiendo la misma cantidad de interés en diferentes unidades.

La organización de esta tesis comprende un primer capítulo en el que se presenta el problema biológico, indicando las características de la enfermedad y las vacunas disponibles para la prevención. En el segundo capítulo se exponen los fundamentos del análisis del modelo de ecuaciones estructurales, la notación empleada, los métodos de estimación de parámetros y las medidas de ajuste del modelo. En el Capítulo 3 se presenta la propuesta del modelo de ecuaciones estructurales para el problema de la evaluación de la calidad de la vacuna contra BoVH-1, definiendo las variables latentes y manifiestas, el modelo de medida y estructural, el diagrama de senderos, y la matriz de covarianza en función de los parámetros del modelo. Asimismo, en este capítulo, se presenta la base de datos empleada. Los resultados de la aplicación del modelo se exponen en el Capítulo 4, y se discuten en el Capítulo 5. Finalmente, la tesis se completa con las conclusiones en el Capítulo 6.

Capítulo 1: EL PROBLEMA BIOLOGICO

1.1. Rinotraqueítis Infecciosa Bovina

La Rinotraqueítis infecciosa bovina (IBR) es una enfermedad respiratoria y reproductiva frecuente del ganado bovino producida por la infección con el herpesvirus bovino tipo 1 (BoHV-1) (OIE, 2010; Thiry *et al.*, 2008). Este es un virus ADN, miembro de la familia herpesviridae, subfamilia alphaherpesviridae, género varicellovirus. Dentro del género se reconocen dos patógenos que afectan a los bovinos: (i) el virus herpes bovino tipo 1 (BoHV-1) asociado a enfermedad respiratoria y reproductora y (ii) el virus herpes bovino tipo 5, (BoHV-5) con capacidad neurovirulenta (Del Medico *et al.*, 2009). Basado en el análisis del ADN con enzimas de restricción se pueden diferenciar varios subtipos de ambos agentes. Aguilar Setién (1987) describe las propiedades biofísicas, bioquímicas y biológicas del virus de la Rinotraqueítis Infecciosa Bovina. El nombre de la enfermedad indica el signo más prominente. Después de un periodo de incubación de 2 a 4 días, comienza la descarga serosa nasal, salivación, fiebre, inapetencia y depresión. En pocos días la descarga nasal y ocular se vuelven mucopurulentas, las lesiones necróticas de la nariz progresan y se vuelven pustulosas y las úlceras se cubren de una pseudo membrana que obstruye las vías superiores y conducen a una respiración bucal. La infección induce al aborto y a la reducción de la producción láctea (Hage *et al.*, 1998). Cuando se practica la monta natural, la lesión genital lleva a la vulvovaginitis pustular infecciosa (VPI) o a la balanopostitis (BP), caracterizadas por lesiones necróticas leves a severas, de la mucosa vaginal o prepucial. Después de la infección por semen, comienza la endometritis (Kendrick y McEntree, 1987). Los terneros infectados con BoHV-1 desarrollan una enfermedad sistémica con lesiones focales necróticas en las vísceras y gastroenteritis. Muchas infecciones cursan subclínicamente (Van Oirschot *et al.*, 1993). La meningoencefalitis aparece como resultado de la infección con una cepa relacionada pero distinta del herpesvirus recientemente propuesta como tipo 5 (Studdert, 1994) aunque el tipo 1 puede causar meningoencefalitis esporádicamente. El tipo 1 afecta a animales de todas las edades pero es más común en animales mayores de 6 meses de edad (OIE, 2004). Los casos de enfermedad respiratoria o genital causada por el tipo 1 que no se complican pueden curar en 5-10 días. La infección bacteriana secundaria, que puede ser producida por bacterias del género *Pasteurella*, puede dar lugar a signos clínicos severos debido a que se afectan las vías respiratorias inferiores.

La transmisión del virus es por contacto directo entre animales o indirecto a través de semen, agua o alimentos contaminados, siendo la puerta de entrada el tracto respiratorio superior, las conjuntivas o el tracto genital. El virus entra en el animal vía nasal y se replica con alto título en las membranas mucosas de las vías respiratorias superiores y las tonsilas diseminándose a la conjuntiva. Por vía del axón neuronal se transporta hasta alcanzar el ganglio trigémino. En la infección genital, el tipo 1 se replica en la membrana mucosa de vagina y prepucio y queda en forma latente en el ganglio sacro. El virus permanece en las neuronas de los ganglios probablemente durante toda la

vida del animal. El estrés provocado por el transporte o el parto, pueden inducir a la reactivación de la infección latente y, en consecuencia, el virus puede diseminarse intermitentemente en el ambiente. La infección da lugar normalmente a una respuesta de anticuerpos y a una respuesta inmune mediada por células, dentro de los 7 a 10 días. Se presume que la respuesta inmune persiste durante toda la vida aunque no pueda detectarse por medio de pruebas usuales. Sin embargo, la inmunidad después de la infección no es de por vida: el ganado puede reinfectarse. Los anticuerpos maternos se transfieren vía calostro a las crías y tienen una vida media de cerca de tres semanas pero pueden detectarse ocasionalmente en animales mayores de 9 meses de edad y raramente en animales mayores. Se sospecha la infección con tipo1 sobre la base de signos clínicos patológicos y epidemiológicos, pero para hacer el diagnóstico definitivo se necesita la ayuda del laboratorio para detectar el virus causal o los componentes virales y la respuesta de anticuerpos específica que ellos inducen (OIE, 2010).

Debido a que la latencia del virus es una secuela normal de la infección con BoHV-1 y la respuesta de anticuerpos (Ac) parece durar de por vida, cualquier animal seropositivo debería considerarse como un portador potencial y un dispersor intermitente del virus, con la excepción de los terneros jóvenes con anticuerpos maternos pasivos y el rodeo no infectado que ha sido vacunado con vacunas a virus muertos o con vacunas marcadoras a virus deletados (OIE, 2010).

1.2. Vacunas

Las únicas herramientas disponibles para controlar la infección en el caso de las infecciones virales, son las medidas de tipo higiénico o la vacunación. Esta última es la forma más eficaz y económica de prevención de enfermedades bovinas. La vacunación contra herpesvirus también está recomendada dada la distribución y alta contagiosidad del virus. La alta seroprevalencia en la región de herpesvirus, que en nuestro país, en el sudoeste bonaerense y sur de Corrientes, se ha estimado en 30 a 60% de animales (Odeón *et al.*, 2001), hace también necesaria una clarificación sobre las posibles medidas de control y el rol de las vacunas. Se ha demostrado que las vacunas no previenen formas de latencia, o reactivación del virus y que los animales vacunados, dependiendo de la capacidad inmunogénica de la vacuna, pueden infectarse luego de la exposición al virus, pero en forma asintomática o desarrollando síntomas muy leves de la enfermedad y pueden excretar virus, aunque a concentraciones más bajas y por menor tiempo, que los no vacunados. Una vacuna de potencia satisfactoria es aquella que reduce significativamente la severidad del cuadro clínico y la excreción viral en al menos dos logaritmos respecto de los controles no vacunados (OIE, 2010).

Odeón (2002) describe los distintos tipos de vacunas. Existen vacunas con virus vivo modificado y con virus muerto, o inactivado. En las primeras, el virus está atenuado por multiplicación en una especie no susceptible o en cultivos en laboratorio. El virus inoculado con la vacuna tiene una capacidad de multiplicación limitada en el organismo, pero suficiente como para generar una respuesta inmune (similar a la que se origina con la enfermedad). El virus pierde su capacidad de causar enfermedad debido a que está atenuado, reduciéndose también su eliminación o contagio por parte de los animales vacunados. Luego de la vacunación con virus vivo modificado la protección se establece en unas 2 a 3 semanas, obteniendo protección generalizada (sistémica) y en mucosas (oral, nasal, genital) por períodos prolongados, de meses y hasta años.

Si bien las vacunas con virus vivo producen protección eficiente y su producción es económica, presentan considerables desventajas tales como el riesgo de abortos (Wyler *et al.*, (1989); Kahrs, (1977); Miller *et al.*, (1991); Mitchell, (1974)), y el desarrollo de infecciones latentes con la cepa del virus (Nettleton *et al.*, (1984). Asimismo, pueden causar un estado temporario de inmunosupresión acompañado de respuesta inmune deficiente y susceptibilidad aumentada a otras infecciones (Yates, 1982). Estas vacunas no están autorizadas en nuestro país.

Las vacunas con virus inactivado no tienen los inconvenientes de las vacunas con virus vivo o atenuado mencionadas. Al utilizar virus muerto, no presentan riesgo biológico, por lo que también pueden aplicarse a hembras gestantes. Sin embargo, debido al proceso de inactivación, las vacunas con virus muerto generan una respuesta inmune limitada (no generan inmunidad en mucosas o inmunidad mediada por células) por lo que, generalmente, se requieren adyuvantes y dosis múltiple para dar una protección adecuada.

Puertes Colorado (2009) refiere que un inconveniente de las vacunas convencionales es que no permiten diferenciar entre animales vacunados e infectados y por tanto interfieren con los sistemas de diagnóstico que se utilizan en las campañas de control y erradicación, pudiendo un país perder su estatus de libre de enfermedad. Esto ha llevado a que, a pesar de reconocerse su eficacia, se prohíba la utilización de vacunas en los programas de control sanitario de determinadas enfermedades como, por ejemplo, la fiebre aftosa o la peste porcina clásica. Los avances experimentados en los últimos años en el dominio de las técnicas de ingeniería genética y en el conocimiento de la biología molecular de los patógenos y los grandes avances que han tenido lugar en el campo de la inmunología, fundamentalmente en el conocimiento de la inmunidad innata y adquirida y de cómo las células T y las células B reconocen sus respectivos determinantes antígenos y de cómo se comunican las distintas poblaciones entre sí, han orientado e impulsado los estudios sobre nuevos sistemas de vacunación más seguros y eficientes. Las nuevas tecnologías en el campo de las vacunas veterinarias han abierto la posibilidad de diseñar vacunas marcadoras en las cuales se ha eliminado el gen codificante para una de las proteínas virales, de manera que el virus vacunal no inducirá respuesta inmune para este antígeno viral a diferencia del virus infeccioso (vacunas que Diferencian animales Infectados de Vacunados, DIVA). Estas vacunas en combinación con un sistema de diagnóstico adecuado, permiten la diferenciación entre animales vacunados sanos y animales infectados. Las vacunas marcadoras son herramientas claves en las campañas de control y erradicación de las enfermedades y facilitan el mercado y movimiento de animales. Con parte de Europa libre de BoVH-1, la habilidad para diferenciar no sólo animales infectados de vacunados sino también animales vacunados que se han infectado, se ha convertido en crítica como herramienta de comercialización y monitoreo (Zhao y Xi, 2011). Las vacunas marcadoras con la proteína gE eliminada, ya sea con virus vivo o inactivado, son ampliamente utilizadas actualmente en Europa en conjunto con las pruebas diagnósticas para monitorear el ganado basadas en gE (Van Drunen Littel-Van den Hurk, 2006).

Más adelante, Puertes Colorado (2009) expresa que la tendencia en salud animal es a la obtención de vacunas perfectamente definidas a nivel molecular que selectivamente activen la respuesta inmunitaria deseada. La vacuna ideal debería activar este tipo de respuesta y establecer una inmunidad protectora, fundamentalmente frente a la infección, de forma rápida y duradera. El establecimiento rápido de la inmunidad es un factor importante en el caso de vacunaciones en situaciones de emergencia donde se persigue frenar la extensión del brote infeccioso.

Como se mencionó, un problema que presentan las vacunas inactivadas, es que son débiles en relación a la inmunogenicidad y necesitan ser combinadas con adyuvantes inmunitarios para ser potenciadas. Los adyuvantes fueron descritos por primera vez por Ramon (1924) como sustancias que, combinadas con un antígeno, producían una respuesta inmunitaria más fuerte que la inducida por el antígeno sólo. Desde entonces se han valorado muchas sustancias naturales y sintéticas para potenciar la respuesta inmunitaria y la eficacia de las vacunas. Los adyuvantes inmunológicos constituyen un grupo amplio y heterogéneo de sustancias capaces de estimular el sistema inmunitario de una manera inespecífica y por tanto aumentar la respuesta contra sustancias débilmente antigénicas.

Continuando con la disertación de Puertes Colorado (2009), el autor expresa que una de las características de la respuesta del sistema inmunitario frente a un estímulo antigénico es su heterogeneidad potencial. El carácter de la respuesta (humoral, celular, local o sistémica) está sujeto al proceso de inmunización y a la naturaleza de la interacción del antígeno con las células del sistema inmunitario. Es en esta interacción donde el vehículo que se utilice para administrar el antígeno juega un rol muy importante. Actualmente se tiende hacia un diseño racional de vacunas en el que se empleen, por un lado, estructuras antigénicas perfectamente caracterizadas y, por otro, sistemas adyuvantes que sean capaces de estimular el tipo de respuesta inmunitaria deseada y que, además, sean tolerados por el individuo. La incorporación de un adyuvante puede hacer disminuir la dosis de antígeno necesaria, aspecto importante cuando la capacidad de producción es limitada. También puede mejorar la estabilidad de la vacuna. El adyuvante que se seleccione puede permitir vías de administración distintas a la vía parenteral.

Los adyuvantes se pueden agrupar por su estructura en adyuvantes *particulados* y *no particulados*. Los adyuvantes particulados (O'Hagan y Singh, 2003) incluyen una serie de sustancias que existen como partículas y deben, en parte, su mecanismo de acción, a esta propiedad. Se trata de sustancias que provocan la formación de un depósito del antígeno en el lugar de inyección, del cual se libera lentamente ampliando el tiempo de interacción entre el antígeno y la célula presentadora. Dentro de este tipo de adyuvantes se encuentran, entre otros:

- Las sales minerales como el fosfato de aluminio o el hidróxido de aluminio. El gel de hidróxido de aluminio por su baja toxicidad es el único adyuvante autorizado en vacunas humanas y es un componente que con frecuencia forma parte de la composición de las vacunas veterinarias. Da lugar a un depósito de corta duración, es un buen vehículo e induce una fuerte respuesta del tipo Th2 (Cox y Coulter, 1997).
- Las emulsiones oleosas del tipo agua/aceite o agua/aceite/agua a base de aceites minerales o de aceites metabolizables, utilizadas comúnmente en vacunas veterinarias. Este tipo de adyuvantes forman depósitos más duraderos y estimulan una buena respuesta humoral. Sin embargo, el aceite puede persistir en el sitio de inyección por un período largo de tiempo, puede causar una excesiva reacción del tejido (Rhee *et al.*, 1987) y presentar carcinogenicidad (Murria *et al.*, 1972; Potter y McCardle, 1964).
- El MF59, que es una emulsión oleosa del tipo aceite/agua compuesta a base de aceite de escualeno. Este adyuvante ha sido autorizado para uso en humanos (Peek *et al.*, 2008) por su buena tolerancia. Actúa como vehículo y favorece la presentación del antígeno a las células inmunocompetentes. El escualeno es un constituyente normal del cuerpo del animal y es considerado biocompatible y biodegradable. El escualeno es un hidrocarburo no saturado y susceptible de

oxidación, y es considerado como no deseado para las aplicaciones farmacológicas. Por esta razón, el escualeno fue elegido como una alternativa al aceite mineral en la formulación del adyuvante completo Freund's. Hilger *et al.* (1994, 1999) describen un adyuvante oleoso no-mineral para vacunas veterinarias.

- El complejo inmunoestimulante ISCOM (Morein *et al.*, 1987), que es una matriz con una estructura estable que resulta de la interacción de una proteína antigénica, el adyuvante Quil A, moléculas de colesterol y fosfolípidos. Este complejo tiene la gran ventaja de permitir la incorporación en una misma estructura del antígeno y del adyuvante. Este tipo de estructuras ha demostrado estimular una potente respuesta inmunitaria tanto humoral como celular.
- Los liposomas, esferas concéntricas compuestas de bicapas de fosfolípidos. Sus propiedades adyuvantes dependen del número de capas que lo forman, la carga neta, la posición lipídica, el tamaño y el modo de localización del antígeno dentro del mismo (Gregoriadis, 1990). Por su naturaleza versátil los liposomas permiten la incorporación de otras moléculas inmunomoduladoras tanto lipófilas como hidrófilas y así consiguen un efecto sinérgico.
- Dentro de los nuevos sistemas adyuvantes particulados se pueden citar las nano y micro partículas constituidas por polímeros biodegradables, fundamentalmente el ácido poli-DL-láctico-co-glucólico (PLG) y el ácido poliláctico (PLA). Este tipo de polímeros son prácticamente inocuos y se utilizan en medicina humana para la fabricación de suturas e implantes. El principal atractivo de estas partículas es que se puede controlar la velocidad de liberación y por tanto formar un depósito del cual el antígeno encapsulado es liberado gradual y lentamente. Gracias a ello se puede conseguir un aumento de la duración de la inmunidad o incluso el desarrollo de vacunas efectivas a dosis única (Jaganathan *et al.*, 2005). Dentro de los adyuvantes no particulados se agrupan sustancias que presentan una actividad inmunomoduladora. Dentro de esta categoría Puertes Colorado (2009) refiere:
- La saponina y sus derivados Quil A y QS 21 (Raiput, *et al.*, 2007). La saponina es un glicósido triterpénico que se extrae de la planta *Quillaria saponaria*. La saponina es otro de los adyuvantes que se utilizan con frecuencia en vacunas veterinarias, solas o combinadas con el hidróxido de aluminio. Su actividad adyuvante es consecuencia de su actividad como agente tensoactivo y de la capacidad que tiene de unirse al colesterol en las membranas celulares formando poros y favoreciendo la presentación del antígeno a las células inmunocompetentes a través de la ruta endógena, permitiendo la estimulación de la respuesta T citotóxica.
- Otro adyuvante con actividad inmunoestimulante es el monofosforil lípido A, un derivado del lipopolisacárido (LPS) de *Salmonella Minnesota*, que conserva la mayoría de las propiedades adyuvantes del LPS pero con una toxicidad reducida. Es un agonista del receptor TLR4 (*toll-like*) y, entre otros efectos, provoca una fuerte respuesta Th1.
- Las citoquinas, que son mediadores endógenos de naturaleza glicoproteica capaces de modular la respuesta inmunitaria. Dentro de ellas se puede citar como ejemplo la interleucina (IL)-1, la IL-2, la IL-12 y el interferón (IFN)- γ . A pesar de su actividad, su costo de producción es elevado y presentan los inconvenientes de ser específicas de especie, tener cierta toxicidad y por su naturaleza proteica ser inestables.

1.3. Calidad de las vacunas

Las vacunas multivalentes fueron designadas para controlar un problema sanitario de etiología compleja y para facilitar el manejo del calendario de vacunación de los rodeos. Sin embargo, la potencia y la eficacia contra cada antígeno presente en estas formulaciones combinadas debe ser evaluada.

Específicamente, para la aprobación de vacunas para BoHV-1 las políticas regulatorias internacionales recomiendan evaluar la calidad de las vacunas en pruebas de potencia en bovinos seronegativos (OIE, 2008; CFR, 1985). Una vacuna para BoHV-1 debe prevenir el desarrollo de síntomas clínicos severos y reducir marcadamente la excreción de virus luego del desafío experimental. En general, la Farmacopea Europea establece que en cada lote de producto final se controle: la esterilidad o pureza, las características generales, la seguridad en una de las especies de destino, la identidad del principio activo y la potencia.

Puertes Colorado (2009) expone que en el proceso de desarrollo de una vacuna, se debe utilizar un lote de medicamento producido según el proceso de fabricación propuesto. Éste se debe administrar por todas las vías de administración recomendadas a animales de cada especie (especie destino y especie alternativa). El autor se refiere, también, a la necesidad de evaluar el riesgo medioambiental y valorar los posibles efectos nocivos derivados del uso del medicamento.

La eficacia de las vacunas veterinarias debe ser demostrada por estudios estadísticos de vacunación y desafío, en el animal hospedador usando el animal más joven por el cual el producto es recomendado. Las pruebas deben ser llevadas a cabo bajo estrictas medidas de control con animales seronegativos. **Si existieran pruebas de potencia serológicas validadas, entonces las pruebas de vacunación-desafío no serían requeridas** (CFR 1985; OIE 2010).

Para la evaluación de la calidad inmunológica de una vacuna se considera la respuesta humoral y celular post-vacunación. La primera se refiere a la detección de anticuerpos contra BoHV-1 en suero mediante pruebas de seroneutralización (SN) (Bitsch, 1978) y varios ELISA (Kramps *et al.*, 1993). En OIE (2010) se presenta el protocolo para la prueba de SN. Los métodos ELISA para detectar anticuerpos contra BoHV1 parecen estar reemplazando gradualmente las pruebas SN. No se ha establecido un procedimiento ELISA estándar. Hay disponibles varios tipos de ELISA, incluyendo ELISAs indirectos y de bloqueo. Los procedimientos indirectos son los de uso más común. Existen diversas preparaciones comerciales ELISA, muchas de las cuales son también adecuadas para detectar anticuerpos en la leche. Por razones de estandarización en un país o estado, sería deseable comparar la calidad de las preparaciones comerciales y probar cada lote respecto a criterios definidos previamente en un laboratorio nacional de referencia, antes de su uso en otros laboratorios del país (OIE, 2010).

La respuesta celular está mediada por distintas células del sistema inmune, fundamentalmente por **linfocitos T y macrófagos** y, en menor medida, células NK, neutrófilos y basófilos. La valoración de la respuesta celular se emplea para conocer el estado inmunitario de los animales y la respuesta inmune a la vacunación de formulaciones con adyuvantes específicos y especialmente en la respuesta a vacunas atenuadas donde la replicación del virus vacuna en el hospedador activa estos mecanismos inmunológicos. (Gómez *et al.*, 2011)).

La eficacia de la vacuna se debe demostrar en un experimento de desafío en condiciones de laboratorio. En este tipo de experimento, se administra la vacuna a terneros seronegativos para BoHV1 manteniendo a algunos terneros como control. Todos los terneros se inoculan intranasalmente 3 semanas después con una cepa

virulenta de BoHV1 que origina los síntomas típicos de la enfermedad. El título máximo de virus en la mucosa nasal de los terneros vacunados debe ser por lo menos 100 veces inferior al de los terneros control. El período de excreción del virus debe ser al menos 3 días más corto en los terneros vacunados que en los terneros control (OIE, 2010).

Las pruebas con bovinos son engorrosas y costosas en dinero y tiempo, especialmente en países como Argentina, donde la enfermedad es altamente endémica (Odeón *et al.*, 2001; Campero *et al.* 2003; Moore *et al.* 2003). La dificultad de encontrar animales seronegativos en este escenario lleva a la necesidad de desarrollar pruebas estandarizadas y armonizadas en animales de laboratorio. La disponibilidad de un modelo animal de laboratorio permite al organismo regulatorio y a los productores de vacunas llevar a cabo pruebas de rutina disminuyendo los costos en tiempo y dinero.

Alonzo *et al.* (2008) evaluaron la inmunogenicidad de vacunas en Herpesvirus del bovino a partir de cobayos. Por otra parte, Parreño *et al.*, 2010a trabajaron en la validación de un método alternativo para la prueba de potencia de vacunas para BoHV-1 en cobayos. La prueba fue especialmente diseñada para evaluar la inmunogenicidad a BoHV-1 y otros antígenos virales dentro del grupo de vacunas combinadas inactivadas (Taffs, 2001). Fue desarrollada en los productos terminados en un modelo *in vivo*, con el objetivo de detectar no sólo la variabilidad en cantidad y calidad del antígeno sino también las variables que afectan la inmunogenicidad, tales como la potencia del adyuvante, la estabilidad, y la variabilidad durante el proceso de formulación (McVey *et al.*, 2003).

Capítulo 2: FUNDAMENTOS DE LOS MODELOS DE ECUACIONES ESTRUCTURALES

2.1. Ecuaciones estructurales

La modelización con ecuaciones estructurales (SEM) abarca una familia entera de modelos conocidos con diversos nombres, entre ellos análisis de la estructura de la covariancia, análisis de variable latente, análisis de factor confirmatorio y, con frecuencia, simplemente análisis LISREL, que es el nombre de uno de los programas de *software* más populares. Bollen (1989) provee un tratamiento amplio de estos modelos.

Todas las técnicas SEM se distinguen por dos características:

- (1) estimación de relaciones de dependencia múltiple y cruzada, y
- (2) capacidad de representar los conceptos no observados en estas relaciones y tener en cuenta el error de medida en el proceso de estimación.

La modelización con ecuaciones estructurales permite acomodar relaciones de dependencia cruzadas múltiples en un único modelo. SEM puede estimar varias ecuaciones a la vez, que pueden estar interrelacionadas, por lo que la variable dependiente en una ecuación puede ser una variable independiente en otra(s) ecuación(es). Esta característica permite modelizar relaciones complejas que no sería posible modelar con ninguna otra técnica multivariada.

SEM también permite incorporar **variables latentes** al análisis. Una variable latente es un concepto supuesto y no observado que puede ser sólo aproximado mediante variables medibles u observables. Las variables observadas se conocen como **variables manifiestas**. La incorporación de variables latentes tiene justificaciones tanto prácticas como teóricas, al mejorar la estimación estadística, mejorar la representación de los conceptos teóricos y tener en cuenta el error de medida.

A diferencia de las técnicas de regresión, en este procedimiento se enfatiza el término “covariancia” en vez de “casos”. En lugar de minimizar funciones de valores individuales de observados y predichos, se minimiza la diferencia entre las covariancias muestrales y las covariancias predichas por el modelo. La hipótesis fundamental para los procedimientos de ecuaciones estructurales es que **la matriz de covariancia de las variables observadas es una función de un conjunto de parámetros**. Si el modelo fuese correcto y si se conocieran los parámetros, la matriz de covariancia poblacional sería reproducida exactamente.

La ecuación que formaliza esta hipótesis fundamental es:

$$\Sigma = \Sigma(\theta) \quad (2.1)$$

donde Σ es la matriz de covariancia poblacional de las variables observadas y θ es el vector que contiene los parámetros. La expresión $\Sigma(\theta)$ indica la matriz de covariancias escrita como una función de los parámetros del modelo.

El término *ecuaciones estructurales lineales* hace referencia a que las relaciones entre las variables pueden ser representadas por ecuaciones de estructura lineal o pueden ser transformadas a formas lineales. Se excluyen las ecuaciones estructurales no lineales

en los parámetros. Sin embargo, las funciones no lineales de parámetros son comunes en la *ecuación estructural de covariancia*, $\Sigma = \Sigma(\theta)$. De modo que las ecuaciones estructurales que relacionan las variables observadas, latentes y de disturbio son lineales, pero no necesariamente lo son las ecuaciones estructurales de covariancia.

El término **estructural** se refiere al supuesto de que los parámetros no son sólo medidas descriptivas de asociación sino que revelan una relación “causal” invariante. La técnica no descubre relaciones causales, sino que, en el mejor de los casos, muestra si los supuestos de causalidad incluidos en el modelo se corresponden con una muestra de datos.

2.2 Relación entre variables latentes y observadas

El proceso de medida comienza con un *concepto*. Un concepto es una idea que unifica fenómenos bajo un único término. En los modelos de medida los conceptos están representados por las variables latentes. La medición es un proceso por el cual un concepto es ligado a una o más variables latentes, y éstas son ligadas a variables observadas. Una o más variables latentes pueden ser necesarias para representar un concepto.

Una vez que el concepto es seleccionado o diseñado, se siguen cuatro pasos en un modelo de medida: (1) dar el significado del concepto, (2) identificar las dimensiones y las variables latentes que lo representan, (3) formar medidas, y (4) especificar la relación entre las **medidas** y las **variables latentes**.

El primer paso se lleva a cabo desarrollando una definición teórica que explica de un modo tan simple y preciso como sea posible el significado de un concepto. Cumple varias funciones útiles. En primer lugar, una definición teórica aparea un término y un concepto detallando el sentido específico dado al término. Segundo, clarifica las dimensiones de un concepto. Las dimensiones son los distintos aspectos de un concepto. Son componentes que no pueden ser fácilmente subdivididas en componentes adicionales. Como muchos conceptos tienen numerosas dimensiones posibles, una definición resulta crítica para poner límites a las dimensiones seleccionadas. Se requiere una variable latente por cada dimensión. En tercer lugar, una definición teórica provee de guía para la selección de medidas.

El siguiente paso en la medición es formar medidas. Este paso depende de la definición teórica que, en esta etapa, es a veces referida como la definición operativa. Esta definición describe los procedimientos a seguir para formar medidas de las variable(s) latente(s) que representan un concepto. Una definición operativa o medida es apropiada en el sentido que da lugar a una variable observada que corresponde al significado asignado a un concepto.

El cuarto paso de la medición es construir el modelo de medida. Un modelo de medida especifica un modelo estructural que conecta variables latentes a una o más medidas o variables observables. El modelo de medida describe la relación entre la medida y las variables latentes. Esta relación puede presentarse en una ecuación o en un diagrama de senderos.

Un problema que se presenta en el modelo de medida es el de la selección de la escala para la variable latente. Algunas variables latentes no tienen una escala intrínseca. Las escalas empleadas emergen del consenso, y para muchos conceptos no existe consenso. En estas situaciones, una solución es asignar una escala poniendo uno

de los coeficientes, λ_{ij} , igual a 1. Esto resulta en una variable latente que tiene la misma escala que la variable observada i -ésima.

Dos propiedades importantes de las medidas son su validez y su fiabilidad. La validez de la medida se refiere a si la medida mide lo que se supone debe medir. A pesar de que nunca se puede probar la validez de la medida, se puede desarrollar una base sólida para ella. La fiabilidad se refiere a la consistencia de la medida. Ambas propiedades no se evalúan en forma dicotómica (se cumple o no se cumple) sino sobre un continuo.

2.3 Componentes del sistema y notación

La notación que se va a emplear en este trabajo es la correspondiente a la desarrollada por Jöreskog (1973, 1977), Wiley (1973) y Keesling (1972) y que es conocida como notación LISREL (*Linear Structural RELationships*), citada por Bollen (1989).

El modelo completo consiste en un **sistema de ecuaciones estructurales**. Las ecuaciones contienen **variables aleatorias**, **parámetros estructurales**, y, a veces, **variables no aleatorias**. Las variables no aleatorias son variables explicativas cuyos valores permanecen constantes en muestreos aleatorios repetidos (variables fijas o no estocásticas). Estas son menos comunes que las variables explicativas aleatorias. Los **tres tipos de variables aleatorias** son: **latentes**, **observadas** y variables de **error o disturbio**.

El nexo entre las variables se resume en los parámetros estructurales. Los parámetros estructurales son constantes invariantes que proveen la relación “causal” entre variables. Los parámetros estructurales pueden describir el nexo causal entre variables no observadas, o entre variables observadas, o entre variables no observadas y observadas.

El sistema de ecuaciones estructurales tiene dos subsistemas fundamentales:

- **el modelo de variable latente**
- **el modelo de medida**

El modelo general de ecuaciones estructurales representa una síntesis de estos dos tipos de modelos. Consiste en un modelo de medida que especifica la relación de las variables observadas con las latentes, y un modelo de variables latentes que muestra la influencia de las variables latentes entre sí.

2.3.1. Modelo de variable latente

Las variables latentes representan conceptos unidimensionales en su forma más pura. Otros términos para estas son “variables no observadas” o “variables no medidas” o “factores”. Las variables observadas o indicadores de una variable latente contienen errores aleatorios o sistemáticos, pero las variables latentes están libres de éstos, ya que las variables latentes corresponden a conceptos; son variables hipotéticas. Los

conceptos y las variables latentes, sin embargo, varían en su grado de abstracción. El modelo de variable latente abarca las ecuaciones estructurales que resumen las relaciones entre variables latentes.

Una **variable latente exógena** se simboliza como ξ (xi). Una variable latente es exógena cuando su causa está afuera del modelo. Una **variable latente endógena** es aquella determinada por variables dentro del modelo. Cada variable latente endógena es representada por η (eta). Las variables latentes endógenas son sólo explicadas parcialmente por el modelo. La componente no explicada es representada por ζ (zeta) que es el disturbio aleatorio en la ecuación.

Los términos exógeno y endógeno son específicos del modelo. Puede ser que una variable exógena en un modelo sea endógena en otro, o, una variable especificada como exógena, sea en realidad influenciada por una variable en el modelo. Se puede escribir el modelo general con variables latentes endógenas y exógenas en forma compacta empleando notación matricial como:

$$\eta = \mathbf{B} \eta + \mathbf{\Gamma} \xi + \zeta \quad (2.2)$$

Las variables en el modelo son desviaciones de su media por lo que no se incluyen términos de intercepción. La variable η es un vector de dimensión $m \times 1$ de variables latentes aleatorias endógenas. El vector ξ es $n \times 1$, y representa las n variables latentes exógenas. En la mayoría de los casos ξ es un vector de variables aleatorias. Los errores en las ecuaciones o disturbios están representados por ζ como un vector $m \times 1$. Un ζ_i está asociado con cada η_i , con i que va de 1 a m . El vector ζ generalmente contiene variables aleatorias. Como en regresión, el disturbio ζ_i incluye variables que influyen en η_i pero no son incluidas en la ecuación de η_i . Se asume que los factores omitidos en la ecuación originan ζ_i , tienen $E(\zeta_i) = 0$ y no están correlacionados con las variables exógenas en ξ . De otro modo, es probable obtener estimadores inconsistentes de los coeficientes.

Las matrices de los coeficientes son \mathbf{B} y $\mathbf{\Gamma}$. La **matriz \mathbf{B}** es una matriz de coeficientes $m \times m$ para las variables latentes endógenas, muestra la influencia de las variables latentes entre sí. Su elemento típico es β_{ij} donde i y j se refieren a la posición de la fila y la columna. El modelo asume que $(\mathbf{I} - \mathbf{B})$ es no-singular y que existe $(\mathbf{I} - \mathbf{B})^{-1}$. La diagonal principal de \mathbf{B} es siempre igual a cero. Esto sirve para eliminar η_i del lado derecho de la ecuación i -ésima para la cual es la variable dependiente, es decir, se asume que una variable no es una causa inmediata e instantánea de sí misma. Un cero en \mathbf{B} también indica la ausencia de un efecto de una de las variables latentes endógenas sobre otra. La matriz $\mathbf{\Gamma}$ es una matriz de coeficientes $m \times n$ para las variables latentes exógenas. Sus elementos son simbolizados como γ_{ij} .

La **matriz de covariancia de las variables latentes exógenas** $n \times n$ es $\mathbf{\Phi}$ (phi) con elementos ϕ_{ij} . Como toda matriz de covariancia, es simétrica. Si las variancias de las variables ξ son iguales a uno, entonces $\mathbf{\Phi}$ es una matriz de correlación. La **matriz de covariancia $m \times m$ de los errores** en las ecuaciones es $\mathbf{\Psi}$ (psi) con elementos ψ_{ij} . Cada elemento de la diagonal principal de $\mathbf{\Psi}$ (ψ_{ii}) es la variancia de la correspondiente variable η_i que no está explicada por las variables explicativas incluidas en la i -ésima ecuación. La **matriz de covariancia de η** es una función de \mathbf{B} , $\mathbf{\Gamma}$, $\mathbf{\Phi}$ y $\mathbf{\Psi}$. No tiene un símbolo especial.

En el Cuadro 2.1 se presenta una síntesis de la notación para modelos que relacionan variables latentes (Bollen, 1989)

Cuadro 2.1 Notación para modelos de variables latentes

Ecuación Estructural para un Modelo de Variable Latente			
$\eta = B\eta + \Gamma\xi + \zeta$			
Supuestos			
$E(\eta) = 0$			
$E(\xi) = 0$			
$E(\zeta) = 0$			
ζ no correlacionada con ξ			
$(I - B)$ no singular			
Símbolo	Nombre	Dimensión	Definición
Variables			
η	eta	$m \times 1$	Variable latente endógena
ξ	xi	$n \times 1$	Variable latente exógena
ζ	zeta	$m \times 1$	Errores latentes en la ecuación
Coeficientes			
B	beta	$m \times m$	Matriz de coeficientes para las variables latentes endógenas
Γ	gamma	$m \times n$	Matriz de coeficientes para las variables latentes exógenas
Matrices de Covariancia			
Φ	phi	$n \times n$	$E(\xi\xi')$ (matriz de covariancia de ξ)
Ψ	psi	$m \times m$	$E(\zeta\zeta')$ (matriz de covariancia de ζ)

2.3.2. Modelo de medida

Como en el modelo de variables latentes, las variables observadas tienen una variedad de denominaciones que incluyen variables manifiestas, indicadores y *proxies*⁴. Una estrategia es emplear indicadores únicos, mientras que otra opción es construir un índice con dos o más variables manifiestas para cada concepto. El análisis empírico se realiza a partir de estos indicadores observados o índices, y se tratan los resultados como pruebas de la relación entre las variables latentes.

El supuesto que subyace en estas estrategias es que las variables observadas están perfectamente correlacionadas (o al menos casi perfectamente) con las variables latentes que ellas miden. En la mayoría de los casos esto no es cierto. El modelo de medida tiene ecuaciones estructurales que representan la conexión entre las variables latentes y las observadas de un modo imperfecto en lugar de determinístico. El modelo de medida puede ser escrito en forma matricial como:

$$\mathbf{x} = \Lambda_x \xi + \delta \quad (2.3)$$

$$\mathbf{y} = \Lambda_y \eta + \varepsilon \quad (2.4)$$

⁴ Del inglés sustituto, representante.

Como en el modelo de variable latente, las variables en el modelo de medida son desviaciones de su media. Como puede observarse, todas las variables manifiestas dependen de las variables latentes.

Las variables aleatorias en \mathbf{x} son indicadores de las variables latentes exógenas (las ξ). Las variables aleatorias en \mathbf{y} son indicadoras de las variables latentes endógenas (las η). En general, \mathbf{x} es $q \times 1$ (donde q es el número de indicadores de ξ) e \mathbf{y} es $p \times 1$ (donde p es el número de indicadores de η).

Las matrices Λ_x y Λ_y contienen parámetros que corresponden a los coeficientes estructurales que relacionan las variables latentes con las manifiestas. Estos coeficientes representan las magnitudes del cambio esperado en la variable observada por una unidad de cambio en la variable latente. La matriz Λ_x es $q \times n$ (donde n es el número de variables latentes exógenas) y Λ_y es $p \times m$ (donde m es el número de variables latentes endógenas). En análisis factorial confirmatorio o en modelos donde $\mathbf{y} = \boldsymbol{\eta}$ o $\mathbf{x} = \boldsymbol{\xi}$ se puede poner doble subíndice λ (λ_{ij}). El subíndice i se refiere a x_i (y_i) y la j se refiere a ξ_j (η_j) que influyen en x_i (y_i). El vector de los errores de medida para \mathbf{x} es $\boldsymbol{\delta}$, y es $q \times 1$. El vector de errores para \mathbf{y} es $\boldsymbol{\varepsilon}$ y es $p \times 1$. Generalmente, $\boldsymbol{\delta}$ y $\boldsymbol{\varepsilon}$ son vectores de variables aleatorias.

Las matrices $\boldsymbol{\Theta}_\delta$ y $\boldsymbol{\Theta}_\varepsilon$ son matrices de covarianza de los errores de medición. La diagonal principal contiene las variancias de error asociados con los indicadores. Los elementos fuera de la diagonal son las covariancias de los errores de medición para los diferentes indicadores. La matriz $\boldsymbol{\Theta}_\delta$ es $q \times q$ y contiene las variancias del error y covariancias para las variables x , y $\boldsymbol{\Theta}_\varepsilon$ es una matriz $p \times p$ que contiene las variancias del error y sus covariancias para las variables y . Los correspondientes errores de medición pueden estar positivamente correlacionados debido a sesgos sistemáticos presentes, por lo que los elementos fuera de la diagonal pueden ser distintos de cero; es decir, que el error de medición de una variable puede correlacionarse con el error de medición de otra variable. Esta es una de las características más importantes de las ecuaciones estructurales con variables latentes que la distingue del enfoque de la regresión estándar. Estos modelos se acercan más a la realidad ya que permiten tener errores de medición en variables observadas. A su vez, las diferencias sistemáticas en escala son introducidas con los coeficientes λ .

En el Cuadro 2.2 se presenta la síntesis de la notación para el modelo de medida (Bollen, 1989).

Cuadro 2.2 Notación para modelos de medida

Ecuación Estructural para un Modelo de Medida			
$\mathbf{x} = \Delta_x \boldsymbol{\xi} + \boldsymbol{\delta}$			
$\mathbf{y} = \Delta_y \boldsymbol{\eta} + \boldsymbol{\varepsilon}$			
Supuestos			
$E(\boldsymbol{\eta}) = \mathbf{0}$, $E(\boldsymbol{\xi}) = \mathbf{0}$, $E(\boldsymbol{\varepsilon}) = \mathbf{0}$, y $E(\boldsymbol{\delta}) = \mathbf{0}$			
$\boldsymbol{\varepsilon}$ no correlacionado con $\boldsymbol{\eta}$, $\boldsymbol{\xi}$, y $\boldsymbol{\delta}$			
$\boldsymbol{\delta}$ no correlacionado con $\boldsymbol{\eta}$, $\boldsymbol{\xi}$, y $\boldsymbol{\varepsilon}$			
Símbolo	Nombre	Dimensión	Definición
Variables			
\mathbf{Y}		$p \times 1$	Indicadora observada de $\boldsymbol{\eta}$
\mathbf{X}		$q \times 1$	Indicadora observada de $\boldsymbol{\xi}$
$\boldsymbol{\varepsilon}$	epsilon	$p \times 1$	Errores de medición para \mathbf{y}

δ	delta	$q \times 1$	Errores de medición para \mathbf{x}
Coefficientes			
Λ_y	lambda \mathbf{y}	$p \times m$	Coefficiente que relaciona a \mathbf{y} con $\boldsymbol{\eta}$
Λ_x	lambda \mathbf{x}	$q \times n$	Coefficiente que relaciona a \mathbf{x} con $\boldsymbol{\xi}$
Matrices de Covarianza			
Θ_ε	Theta epsilon	$p \times p$	$E(\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}')$ (matriz de covarianza de $\boldsymbol{\varepsilon}$)
$\Theta\delta$	Theta delta	$q \times q$	$E(\boldsymbol{\delta} \boldsymbol{\delta}')$ (matriz de covarianza de $\boldsymbol{\delta}$)

2.4 El modelo general

Para emplear el modelo general se debe especificar el patrón de elementos de cada una de las ocho matrices – \mathbf{B} , $\boldsymbol{\Gamma}$, Λ_y , Λ_x , $\boldsymbol{\Phi}$, $\boldsymbol{\Psi}$, Θ_ε , y $\Theta\delta$. La especificación de estas matrices debe basarse en el conocimiento que se tiene en el área.

Como se dijo anteriormente, la hipótesis básica del modelo general de ecuaciones estructurales está dada por la ecuación (2.1), en donde $\boldsymbol{\Sigma}$ es la matriz de covarianza poblacional de \mathbf{y} y \mathbf{x} , y $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ es la matriz de covarianza escrita como una función de los parámetros del modelo en $\boldsymbol{\theta}$.

La ecuación (2.1) implica que cada elemento de la matriz de covarianza es una función de uno o más parámetros del modelo. Se puede considerar a $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ como un ensamble de tres partes: (1) la matriz de covarianza de \mathbf{y} , (2) la matriz de covarianza de \mathbf{x} con \mathbf{y} , y (3) la matriz de covarianza de \mathbf{x} .

Considerando primero $\boldsymbol{\Sigma}_{yy}(\boldsymbol{\theta})$, la matriz de covarianza de \mathbf{y} :

$$\begin{aligned}
 \boldsymbol{\Sigma}_{yy}(\boldsymbol{\theta}) &= E(\mathbf{y}\mathbf{y}') && \text{(reemplazando de acuerdo al modelo 2.4)} \\
 &= E[(\Lambda_y \boldsymbol{\eta} + \boldsymbol{\varepsilon})(\boldsymbol{\eta}' \Lambda_y' + \boldsymbol{\varepsilon}')] \\
 &= \Lambda_y E(\boldsymbol{\eta} \boldsymbol{\eta}') \Lambda_y' + \Theta_\varepsilon
 \end{aligned} \tag{2.5}$$

Se puede particionar $E(\boldsymbol{\eta} \boldsymbol{\eta}')$, a su vez, sustituyendo la forma reducida de (2.2), esto es,

$$\boldsymbol{\eta} = (\mathbf{I} - \mathbf{B})^{-1}(\boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}), \text{ en (2.5) y simplificando}$$

$$\boldsymbol{\Sigma}_{yy}(\boldsymbol{\theta}) = \Lambda_y(\mathbf{I} - \mathbf{B})^{-1}(\boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Gamma}' + \boldsymbol{\Psi})[(\mathbf{I} - \mathbf{B})^{-1}]' \Lambda_y' + \Theta_\varepsilon \tag{2.6}$$

La matriz de covarianza de \mathbf{y} resulta así en una función complicada de seis de las ocho matrices de parámetros.

La matriz de covarianza de \mathbf{y} con \mathbf{x} es $\boldsymbol{\Sigma}_{yx}$, y al referirse a una función de los parámetros estructurales es $\boldsymbol{\Sigma}_{yx}(\boldsymbol{\theta})$. Se puede escribir

$$\begin{aligned}
 \boldsymbol{\Sigma}_{yx}(\boldsymbol{\theta}) &= E(\mathbf{y}\mathbf{x}') \\
 &= E[(\Lambda_y \boldsymbol{\eta} + \boldsymbol{\varepsilon})(\boldsymbol{\xi}' \Lambda_x' + \boldsymbol{\delta}')] \\
 &= \Lambda_y E(\boldsymbol{\eta} \boldsymbol{\xi}') \Lambda_x'
 \end{aligned} \tag{2.7}$$

Nuevamente, haciendo uso de la forma reducida de $\boldsymbol{\eta}$ sigue que

$$\Sigma_{yx}(\theta) = \Lambda_y(\mathbf{I} - \mathbf{B})^{-1}(\Gamma\Phi\Lambda_x')$$
(2.8)

Finalmente, la matriz de covariancia de \mathbf{x} , Σ_{xx} , escrita como una función de los parámetros estructurales, $\Sigma_{xx}(\theta)$ es:

$$\begin{aligned}\Sigma_{xx}(\theta) &= E(\mathbf{xx}') \\ &= E[(\Lambda_x \xi + \delta)(\xi' \Lambda_x' + \delta')] \\ &= \Lambda_x E(\xi \xi') \Lambda_x' + \Theta_\delta \\ &= \Lambda_x \Phi \Lambda_x' + \Theta_\delta\end{aligned}$$
(2.9)

Si se ensamblan (2.6), (2.8) y (2.9) en una sola matriz $\Sigma(\theta)$, se obtiene la matriz de covariancia para las variables observadas \mathbf{y} , y \mathbf{x} como una función de los parámetros del modelo:

$$\begin{aligned}\Sigma(\theta) &= \begin{bmatrix} \Sigma_{yy}(\theta) & \Sigma_{yx}(\theta) \\ \Sigma_{xy}(\theta) & \Sigma_{xx}(\theta) \end{bmatrix} \\ &= \begin{bmatrix} \Lambda_y(\mathbf{I} - \mathbf{B})^{-1}(\Gamma\Phi\Gamma' + \Psi)[(\mathbf{I} - \mathbf{B})^{-1}]' \Lambda_y' + \Theta_\epsilon & \Lambda_y(\mathbf{I} - \mathbf{B})^{-1}\Gamma\Phi\Lambda_x' \\ \Lambda_x\Phi\Gamma'[(\mathbf{I} - \mathbf{B})^{-1}]' \Lambda_y' & \Lambda_x\Phi\Lambda_x' + \Theta_\delta \end{bmatrix}\end{aligned}$$
(2.10)

La matriz $\Sigma_{xy}(\theta)$ en el cuadrante inferior izquierdo en (2.10) es la transpuesta de $\Sigma_{yx}(\theta)$

2.5 Análisis de Senderos

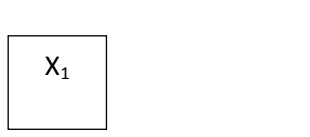

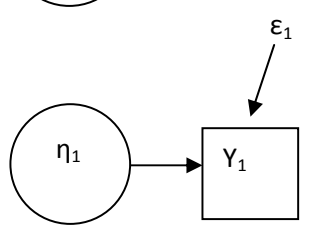

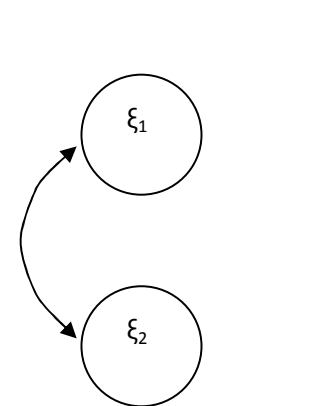
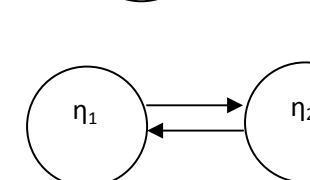
El análisis de senderos, conocido en inglés como *path analysis*, es una metodología para analizar sistemas de ecuaciones estructurales. Las aplicaciones contemplan tres componentes: (1) el diagrama de senderos (*path diagram*), (2) descomposición de covariancias y correlaciones en términos de los parámetros del modelo, y (3) la distinción entre efectos directos, indirectos y totales de una variable sobre otra.

2.5.1 Diagrama de senderos

Un diagrama de senderos es una representación pictórica de un sistema de ecuaciones simultáneas. Una de las principales ventajas de un diagrama de senderos es que presenta una figura de las relaciones que se asumen que existen. En el Cuadro 2.3 se presentan los símbolos principales del diagrama. Las variables observadas están dentro

de cajas rectangulares o cuadradas. Las variables no observadas o latentes están en círculos, con excepción de los términos de disturbio que no están contenidos. Las flechas rectas de una sola punta representan relaciones causales entre las variables conectadas por la flecha. Una flecha curva de dos puntas indica asociación entre dos variables. Las variables pueden estar asociadas por cualquiera de muchas razones. La asociación puede ser debida a que ambas variables dependan de una/s tercera/s, o las variables pueden tener una relación causal pero esta permanece no especificada. La ausencia de flechas conectando variables implica que las covariancias son cero.

Cuadro 2.3 Símbolos principales empleados en el Análisis de senderos

	caja rectangular o cuadrada indica una variable observada o manifiesta
	círculo o elipse indica una variable no observada o latente
	variable no enmarcada indica término de disturbio (error ya sea en la ecuación o la medición).
	flecha recta indica el supuesto de que la variable en la base de la flecha “causa” la variable en la punta de la flecha
	flecha curvada de dos puntas indica asociación entre dos variables
	dos flechas rectas de una punta conectando dos variables indican relación recíproca o causa recíproca

2.6 Identificación del modelo

La identificación es un tema relevante a todos los modelos de ecuaciones estructurales. Los parámetros cuyo estado de identificación es desconocido están en θ , donde θ contiene los t parámetros de \mathbf{B} , $\mathbf{\Gamma}$, $\mathbf{\Phi}$ y $\mathbf{\Psi}$. La ecuación que relaciona Σ con θ es la hipótesis de estructura de covarianza $\Sigma = \Sigma(\theta)$. A modo de definición podríamos decir que si un parámetro desconocido en θ puede ser escrito como una función de uno o más elementos de Σ , ese parámetro es identificado. Si todos los parámetros en θ están identificados, entonces el modelo está identificado.

Una definición alternativa de identificación comienza por considerar vectores θ_1 y θ_2 , cada uno de dimensión $t \times 1$, conteniendo valores específicos para los parámetros desconocidos en θ . Los parámetros en θ están globalmente identificados si no existen vectores θ_1 y θ_2 tales que $\Sigma(\theta_1) = \Sigma(\theta_2)$, a menos que $\theta_1 = \theta_2$.

Un modo de establecer la identificación del modelo es la algebraica, en la que se observa si cada elemento de θ puede resolverse en términos de uno o más elementos conocidos para ser identificados de Σ . El problema de establecer la identificación con este enfoque es que, en modelos complejos, se vuelve una forma muy tediosa y puede llevar a errores.

Existen reglas que pueden ser empleadas para probar si el modelo está globalmente identificado. Una es la *regla t* y otra es la *regla de dos pasos*.

2.6.1 Regla t

La regla t es una condición necesaria pero no suficiente de identificación del modelo, la que se establece si:

$$t < (\frac{1}{2})(p + q)(p + q + 1) \quad (2.11)$$

donde t es el número de elementos libres y no restringidos en θ , entonces el modelo está identificado. La justificación de esta regla es que, si el número de incógnitas en θ excede el número de ecuaciones, la identificación no es posible.

2.6.2 Otras técnicas

Bollen (1989) describe una serie de condiciones necesarias y/o suficientes para la identificación del modelo. Entre ellas la regla de dos pasos. El problema en establecer la identificación de esta forma es que se debe asumir que la matriz Θ_δ es diagonal, supuesto que puede no ser realista.

Existen reglas del tipo “reglas del pulgar” que incluyen, para los modelos de medida, por ejemplo, la “regla de las tres medidas” que establece que un modelo de medida va a estar identificado si cada variable latente está asociada al menos a tres variables indicadoras. La “regla de las dos medidas” establece que un modelo de medida estará identificado si cada variable latente está asociada al menos con dos medidas y, a

su vez, cada variable latente está correlacionada con al menos otra variable latente. Davis (1993) y Reilly (1995) han hecho contribuciones en este aspecto.

En muchas situaciones se necesita recurrir a pruebas empíricas de identificación que son aplicables a todos los modelos de ecuaciones estructurales. Se deben distinguir dos tipos de identificación: (i) la identificación *global* y, (ii) la identificación *local*. Los procedimientos empíricos prueban la identificación local. La definición dada al comienzo de la sección (2.6) es para la identificación global: un vector de parámetros θ está globalmente identificado si no existen vectores θ_1 y θ_2 tales que $\Sigma(\theta_1) = \Sigma(\theta_2)$, a menos que $\theta_1 = \theta_2$. La identificación local es un concepto más débil de unicidad: un vector de parámetros θ está localmente identificado en un punto θ_1 , si en el entorno de θ_1 no existe ningún vector θ_2 para el cual $\Sigma(\theta_1) = \Sigma(\theta_2)$, excepto que $\theta_1 = \theta_2$. Es decir que la identificación local evalúa la identificación en un punto específico de θ , θ_1 , y determina si la matriz de covariancia implícita cambia con pequeños cambios de θ_1 . La identificación global no es tan limitada. Se puede considerar cualquier par de vectores con valores posibles de θ . Si cualquiera de ellos tiene la misma matriz de covariancia implícita, entonces θ no está identificado globalmente. Sin embargo, aún puede estar identificado localmente.

Las pruebas empíricas de identificación están disponibles en programas de estimación de estructuras de covariancia. Una prueba recomendada por Keesling (1972), Wiley (1973), Jöreskog y Sörbom (1986), emplea la matriz de información para θ para determinar la identificación. La matriz de información está basada en los valores esperados de las derivadas parciales segundas de las funciones de ajuste (matriz hessiana), con respecto a cada parámetro en θ :

$$I_n(\theta)_{i,j} = -E \left[\frac{\partial^2 \ln F}{\partial \theta_i \partial \theta_j} \right] \quad (2.12)$$

El parámetro θ está identificado localmente en algún punto θ_1 si, y solo si, existe la inversa de la matriz de información (Rothenberg, 1971). La inversa de la matriz de información estimada es calculada por programas como el SAS, ya que es la base para calcular los errores estándar de $\hat{\theta}$. Así, la falta de una inversa significa que θ en $\hat{\theta}$ no está identificado localmente.

Existen dos problemas con esta prueba. Uno es que para evaluar la existencia de una inversa se apoya en métodos numéricos, y éstos tienen un límite de exactitud. Puede ocurrir que, a causa de un truncamiento numérico, se pueda calcular la inversa, siendo que ésta no existiría si los valores que entran en el cálculo se registraran con un nivel mayor de precisión. Un segundo problema es que, el objetivo es evaluar identificación en θ , mientras que la prueba empírica es para $\hat{\theta}$.

Bollen (1989) citando a Jöreskog y Sörbom (1986), sugiere que si existen aún dudas de la identificación del modelo, como primer paso, se analice la matriz de covariancia observada, S , como de costumbre, y que se guarde $\hat{\Sigma}$, donde $\hat{\Sigma}$ es la matriz de covariancia estimada basada en las estimaciones de los parámetros del modelo. Luego se sustituya $\hat{\Sigma}$ por S y se vuelva a correr el mismo programa. Si el modelo está identificado, las nuevas estimaciones deberían ser idénticas a las primeras que generaron $\hat{\Sigma}$. Otra prueba empírica es estimar el modelo con valores iniciales diferentes para los parámetros y observar si converge a las mismas estimaciones cada vez.

2.7 Estimación de los parámetros del modelo

2.7.1 Métodos de estimación

Los procedimientos de estimación derivan de la relación de la matriz de covariancia de las variables observadas con los parámetros estructurales. Siendo \mathbf{S} la matriz de covariancia muestral para las variables observadas y $\hat{\Sigma}$ la correspondiente a la matriz de covariancia implícita, $\Sigma(\theta)$, se eligen los valores de $\hat{\theta}$ tales que $\hat{\Sigma}$ esté lo más cerca posible de \mathbf{S} . Para definir “qué tan cercanos” están, se necesita una función que sea minimizada. Existen muchas funciones diferentes para esta tarea. Las funciones de ajuste $F(\mathbf{S}, \Sigma(\theta))$ se basan en \mathbf{S} , la matriz de covariancia muestral y en $\Sigma(\theta)$, la matriz de covariancia implícita de los parámetros estructurales. Las estimaciones de θ se sustituyen en $\Sigma(\theta)$. El valor de la función ajustada para $\hat{\theta}$ es $F(\mathbf{S}, \hat{\Sigma})$. Las funciones de ajuste que se presentan aquí tienen las siguientes propiedades:

- (1) $F(\mathbf{S}, \Sigma(\theta))$ es un escalar
- (2) $F(\mathbf{S}, \Sigma(\theta)) \geq 0$
- (3) $F(\mathbf{S}, \Sigma(\theta)) = 0$ si y solo si $\Sigma(\theta) = \mathbf{S}$
- (4) $F(\mathbf{S}, \Sigma(\theta))$ es continua en \mathbf{S} y $\Sigma(\theta)$.

De acuerdo a Browne (1984) minimizar funciones que cumplen con estas condiciones produce estimadores consistentes de θ . Entre estas funciones se pueden mencionar:

- Mínimos cuadrados no ponderados (ULS)
- Mínimos cuadrados generalizados (GLS)
- Máxima verosimilitud (ML)

En cada caso, el vector de los parámetros es estimado iterativamente por un algoritmo no-lineal de optimización que optimiza una función de bondad de ajuste F .

La función de ajuste que es minimizada es, según el método:

a) Mínimos cuadrados no ponderados

$$F_{ULS} = 0.5 \text{ tr } [(\mathbf{S} - \Sigma(\theta))^2] \quad (2.13)$$

Como puede observarse, F_{ULS} , minimiza 1/2 de la suma de cuadrados de cada elemento en la matriz de residuales, $(\mathbf{S} - \Sigma(\theta))$, dándole la misma ponderación a cada elemento, como si los errores en cada ecuación tuvieran la misma variancia y covariancia.

b) Mínimos cuadrados generalizados

$$F_{GLS} = 0.5 \text{ tr } [(\mathbf{S}^{-1}(\mathbf{S} - \Sigma(\theta)))^2] \quad (2.14)$$

A diferencia de la función anterior, en este caso, cada elemento de $(\mathbf{S} - \mathbf{\Sigma}(\boldsymbol{\theta}))$ es ponderado de acuerdo a las variancias y covariancias de cada error residual. Dos supuestos referidos a \mathbf{S} deben cumplirse para que los estimadores tengan óptimas propiedades: (1) $E(s_{ij}) = \sigma_{ij}$, y (2) la distribución asintótica de \mathbf{S} es Normal multivariada con medias σ_{ij} y covariancias asintóticas de s_{ij} y s_{gh} igual a $\mathbf{N}^{-1}(\sigma_{ig} \sigma_{jh} + \sigma_{ih} \sigma_{jg})$.

El supuesto de distribución Normal multivariada asintótica es bastante general. Una condición suficiente para este supuesto es que las observaciones sean independientes e idénticamente distribuidas y que existan los momentos de cuarto orden de \mathbf{x} e \mathbf{y} . El segundo supuesto es satisfecho si \mathbf{x} e \mathbf{y} tienen distribución Normal multivariada pero también se cumple para otras distribuciones sin kurtosis excesiva (Browne, 1982)

c) Máxima verosimilitud basado en la teoría normal

$$F_{ML} = \text{tr}(\mathbf{S}\mathbf{\Sigma}(\boldsymbol{\theta})^{-1}) - (p + q) + \ln|\mathbf{\Sigma}(\boldsymbol{\theta})| - \ln|\mathbf{S}| \quad (2.15)$$

Los estimadores por máxima verosimilitud tienen muchas propiedades importantes. Las propiedades son asintóticas, por lo que se sostienen en muestras grandes. Si bien pueden ser sesgados en muestras pequeñas, los estimadores máximo-verosímiles son asintóticamente insesgados. Además, estos estimadores son consistentes ($\lim_{n \rightarrow \infty} p\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}$). Una tercera propiedad es que son asintóticamente eficientes, es decir que entre los estimadores consistentes, ninguno tiene una variancia asintótica tan pequeña. Asimismo, la distribución de los estimadores se aproxima a la distribución Normal a medida que aumenta el tamaño de la muestra (están asintóticamente normalmente distribuidos). Esta última propiedad permite suponer que, si se conoce el error estándar de los parámetros estimados, el cociente del parámetro estimado y su error estándar debería aproximar una distribución Z en muestras grandes.

Estas tres funciones pueden ser expresadas por un criterio general de mínimos cuadrados ponderados (Browne, 1982):

$$F_{GWLs} = 0.5 \text{tr}[(\mathbf{W}^{-1}(\mathbf{S} - \mathbf{\Sigma}(\boldsymbol{\theta})))^2] \quad (2.16)$$

Para el método de mínimos cuadrados no ponderados, la matriz de ponderación, \mathbf{W} , es la matriz identidad, \mathbf{I} ; para el método de mínimos cuadrados generalizados, la matriz de ponderación es la matriz de covariancia muestral; y para el método de máxima verosimilitud, \mathbf{W} es la matriz de covariancia predicha, $\mathbf{\Sigma}(\hat{\boldsymbol{\theta}})$, actualizada por las iteraciones. Los valores de las funciones de máxima verosimilitud, F_{ML} , y del criterio general de mínimos cuadrados ponderados, F_{GWLs} , con $\mathbf{W} = \mathbf{\Sigma}(\hat{\boldsymbol{\theta}})$ son asintóticamente equivalentes.

Una condición necesaria para la minimización de estas funciones es que las estimaciones de los parámetros sean seleccionadas de modo de que las derivadas parciales de la función de ajuste con respecto a cada estimador sean cero ($\frac{\partial F}{\partial \hat{\boldsymbol{\theta}}} = 0$). Una condición suficiente para que estos valores minimicen F es que la matriz formada

sacando la derivada segunda parcial de la función de ajuste con respecto a cada parámetro estimado (matriz hessiana) sea definida positiva.

Una característica importante que comparten F_{ML} y F_{GLS} , pero no F_{ULS} es que son invariantes en la escala y son “escala-libres” (Swaminathan y Algina, 1978; Coenders *et al.*, 1995). Estas propiedades están referidas a las consecuencias del cambio de las unidades de medida de una o más variables observadas. Una función de ajuste, $F(\mathbf{S}, \mathbf{\Sigma}(\boldsymbol{\theta}))$, es invariante en escala si $F(\mathbf{S}, \mathbf{\Sigma}(\boldsymbol{\theta})) = F[\mathbf{DSD}, \mathbf{D}\mathbf{\Sigma}(\boldsymbol{\theta})\mathbf{D}]$, donde \mathbf{D} es una matriz diagonal no singular con elementos positivos en la diagonal. Por ejemplo, si la diagonal principal de \mathbf{D} consiste en las inversas de los desvíos estándares de las variables observadas, entonces \mathbf{DSD} es la matriz de correlación. La propiedad de ser invariante en escala implica que el valor de la función de ajuste es el mismo cuando \mathbf{DSD} es sustituido por \mathbf{S} y $\mathbf{D}\mathbf{\Sigma}(\boldsymbol{\theta})\mathbf{D}$ es sustituido por $\mathbf{\Sigma}(\boldsymbol{\theta})$. Por esta razón los valores de las funciones de ajuste son los mismos, tanto para la matriz de correlación como de covariancia.

Un concepto relacionado con el anterior es el de escala-libre. Este concepto está referido a que se mantiene la equivalencia entre los parámetros estructurales y los estimadores en un modelo con las variables originales y en un modelo con variables linealmente transformadas. Esto significa que es posible moverse de los datos transformados y no transformados fácilmente conociendo las matrices diagonales de escala.

Otro aspecto importante de los estimadores F_{ML} y F_{GLS} es que proveen una prueba del ajuste global del modelo ya que la distribución asintótica tanto de $(N - 1)F_{ML}$ como de $(N - 1)F_{GLS}$ es una distribución χ^2 con $\frac{1}{2}(p + q)(p + q + 1) - t$ grados de libertad, donde t es el número de parámetros libres y F_{ML} y F_{GLS} son las funciones de ajuste evaluadas en las estimaciones finales de máxima verosimilitud y mínimos cuadrados generalizados, respectivamente. La hipótesis nula de la prueba de χ^2 es $H_0: \mathbf{\Sigma} = \mathbf{\Sigma}(\boldsymbol{\theta})$. Esta hipótesis implica que las restricciones del modelo son correctas. El rechazo de la hipótesis nula sugiere que, al menos una de las restricciones es errónea de modo que $\mathbf{\Sigma} \neq \mathbf{\Sigma}(\boldsymbol{\theta})$.

2.7.2 Soluciones numéricas para minimizar las funciones de ajuste

Como se mencionó en la sección anterior, una condición necesaria para encontrar el mínimo de cualquier función, $F(\boldsymbol{\theta})$, es llevar la derivada parcial de $F(\boldsymbol{\theta})$ a cero con respecto a cada θ_i y resolver para θ_i . Una condición suficiente para que estos valores minimicen a $f(\boldsymbol{\theta})$ es que la matriz de las derivadas segundas parciales (matriz hessiana) sea definida positiva en esos valores.

Si $\boldsymbol{\theta}$ es $t \times 1$ considerando las derivadas parciales de primer orden se tiene que

$$\frac{\partial f(\boldsymbol{\theta})}{\partial \theta_i} = 0, \quad \text{para } i = 1, 2, \dots, t \quad (2.17)$$

Luego de sacar las derivadas parciales, por (2.17) se muestra que hay tantas ecuaciones como θ_i en $\boldsymbol{\theta}$. En el modelo general de ecuaciones estructurales $f(\boldsymbol{\theta})$ es la ecuación de ajuste F_{ML} , F_{GLS} o F_{ULS} , y (2.17) resulta en t ecuaciones que son típicamente no lineales en los parámetros. Generalmente, no son posibles soluciones explícitas para los parámetros y, en estos casos, se hace necesaria la minimización por métodos numéricos.

Los métodos numéricos comienzan con la función objetivo a ser minimizada. Se busca desarrollar una secuencia de valores de θ tal que el último de los vectores en la secuencia minimice una de estas funciones objetivo. El primer valor de prueba se simboliza por $\theta^{(1)}$, el segundo por $\theta^{(2)}$, ..., $\theta^{(j)}$.

Los tres aspectos claves en la minimización son: (1) la selección de los valores iniciales, $\theta^{(1)}$, (2) las reglas para moverse de un paso de la secuencia hacia el próximo (de θ^i a θ^{i+1}), y (3) cuándo finalizar las iteraciones.

2.7.2.1 Valores iniciales

Los valores iniciales en θ , simbolizados como $\theta^{(1)}$, afectan la minimización numérica de distintas formas. Por un lado, determinan el número de iteraciones necesarias para llegar a la solución final ya que valores iniciales cercanos a los finales reducen el número de iteraciones necesarias. Por otra parte, valores iniciales muy alejados de los valores finales aumentan la probabilidad de encontrar mínimos locales en lugar de globales, o de no encontrar una solución convergente.

Existen varias estrategias para seleccionar los valores iniciales. Una es emplear un método no iterativo para estimar los parámetros del modelo y emplear estas estimaciones. Entre estos procedimientos se encuentran la estimación por Variables Instrumentales y Mínimos Cuadrados en dos etapas (2SLS). Los estimadores por variables instrumentales (IV) son no iterativos y proveen estimadores consistentes de los parámetros. Al ser no iterativo, este método es menos costoso en cómputos que la estimación por ML, GLS, o ULS. Además de que es un método de estimación en sí mismo, frecuentemente los estimadores IV proveen de buenos valores iniciales en el proceso iterativo.

El método de estimación IV emplea variables instrumentales para desarrollar un estimador consistente de los parámetros en la ecuación. Estas variables, también conocidas como *instrumentos*, son variables observadas que no están correlacionadas con los disturbios en la ecuación.

El empleo de VI en el área de la Econometría tiene una larga historia en ecuaciones simultáneas en las cuales se trata a las variables observadas como medidas sin error de medición. Bollen (1989, 1996a) extendió la técnica al modelo factorial confirmatorio con errores de medida correlacionados y, posteriormente, desarrolló 2SLS para modelos de medida y de variable latente en SEM, con o sin errores correlacionados a través de algunas ecuaciones.

El estimador por Mínimos Cuadrados en dos etapas (2SLS) presentado por Bollen (1996a) estima términos de intercepción y permite errores correlacionados, y ha sido desarrollado para aplicaciones con disturbios heteroscedásticos (Bollen, 1996b). El procedimiento 2SLS no requiere que las variables observadas provengan de una distribución Normal, y es menos sensible a errores de especificación que las técnicas de información completa como Máxima Verosimilitud.

El método de estimación por IV emplea variables instrumentales para encontrar estimadores consistentes de los coeficientes para las variables explicativas originales y la técnica de 2SLS es un método IV que forma la combinación óptima de IV cuando se puede disponer de más de una IV.

Bollen (1996a) comienza con el modelo de variable latente:

$$\eta = \alpha + B\eta + \Gamma\xi + \zeta \quad (2.18)$$

donde $\boldsymbol{\eta}$ es el vector ($m \times 1$) de las variables latentes endógenas, \mathbf{B} es la matriz ($m \times m$) de coeficientes que describen el impacto de las $\boldsymbol{\eta}$'s entre sí, $\boldsymbol{\xi}$ es un vector ($n \times 1$) de variables latentes exógenas, $\boldsymbol{\Gamma}$ es una matriz ($m \times n$) de coeficientes que describen el impacto de las $\boldsymbol{\xi}$'s sobre las $\boldsymbol{\eta}$'s, $\boldsymbol{\alpha}$ es un vector ($m \times 1$) de términos de intercepción, y $\boldsymbol{\zeta}$ es un vector ($m \times 1$) de disturbios aleatorios con $E(\boldsymbol{\zeta}) = \mathbf{0}$ y $\text{Cov}(\boldsymbol{\xi}, \boldsymbol{\zeta}') = \mathbf{0}$.

El objetivo es estimar $\boldsymbol{\alpha}$, \mathbf{B} y $\boldsymbol{\Gamma}$ del modelo de variable latente, pero para hacerlo es importante, también, conocer el modelo de medida. Dos ecuaciones resumen este aspecto de SEM:

$$\mathbf{x} = \boldsymbol{\tau}_x + \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta} \quad (2.19)$$

$$\mathbf{y} = \boldsymbol{\tau}_y + \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon} \quad (2.20)$$

donde $\boldsymbol{\tau}_x$ es un vector ($q \times 1$) y $\boldsymbol{\tau}_y$ es el vector ($p \times 1$) de términos de intercepción. $E(\boldsymbol{\delta}) = \mathbf{0}$ y la $\text{Cov}(\boldsymbol{\xi}, \boldsymbol{\delta}') = \mathbf{0}$, y $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ y la $\text{Cov}(\boldsymbol{\eta}, \boldsymbol{\varepsilon}') = \mathbf{0}$. Como supuesto adicional es que los $\boldsymbol{\varepsilon}$, $\boldsymbol{\delta}$ y $\boldsymbol{\zeta}$ no están mutuamente correlacionados. Para identificar un modelo de variable latente se deben poner escalas. Una opción sencilla para dar una escala es seleccionar un indicador por cada variable latente para la cual su carga de factor es fijada en 1 y su intercepción en 0. Se asume que la variable de escala está influenciada sólo por una única variable latente y un término de error.

Luego de dar escala a todas las variables latentes, se ordenan los vectores \mathbf{y} y \mathbf{x} :

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \quad \text{y} \quad \mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \quad (2.21)$$

donde \mathbf{y}_1 es el vector ($m \times 1$) de y 's que dan la escala a $\boldsymbol{\eta}$, \mathbf{y}_2 consiste en el vector $((p - m) \times 1)$ de las variables y remanentes, \mathbf{x}_1 es el vector ($n \times 1$) de las x 's que dan la escala a $\boldsymbol{\xi}$, \mathbf{x}_2 es el vector $((q - n) \times 1)$ de las x 's remanentes.

Esto significa que

$$\mathbf{y}_1 = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_1, \quad \text{o} \quad (2.22)$$

$$\boldsymbol{\eta} = \mathbf{y}_1 - \boldsymbol{\varepsilon}_1, \quad (2.23)$$

y

$$\mathbf{x}_1 = \boldsymbol{\xi} + \boldsymbol{\delta}_1, \quad \text{o} \quad (2.24)$$

$$\boldsymbol{\xi} = \mathbf{x}_1 - \boldsymbol{\delta}_1, \quad (2.25)$$

donde $\boldsymbol{\varepsilon}_1$ y $\boldsymbol{\delta}_1$ contienen los errores que corresponden a \mathbf{x}_1 y \mathbf{y}_1 , respectivamente. Sustituyendo (2.23) y (2.25) en (2.18) se llega a

$$\mathbf{y}_1 = \boldsymbol{\alpha} + \mathbf{B}\mathbf{y}_1 + \boldsymbol{\Gamma}\mathbf{x}_1 + \mathbf{u}, \quad (2.26)$$

donde $\mathbf{u} = \boldsymbol{\varepsilon}_1 - \mathbf{B}\boldsymbol{\varepsilon}_1 - \boldsymbol{\Gamma}\boldsymbol{\delta}_1 + \boldsymbol{\zeta}$. Se debe notar que estas manipulaciones algebraicas permiten reformular el modelo de variables latentes en un modelo de ecuaciones simultáneas en el que todas las variables son observadas con excepción del término de disturbio compuesto. Una diferencia importante de (2.26) con un modelo de ecuaciones simultáneas es que no se puede asumir que \mathbf{u} y \mathbf{x}_1 no estén correlacionadas. Se debe tener presente que \mathbf{u} contiene a $\boldsymbol{\delta}_1$ y que $\boldsymbol{\delta}_1$ se correlaciona con aquellas x 's que están medidas con error.

Si se considera una única ecuación, esto es la ecuación i -ésima de \mathbf{y}_1 de (2.26):

$$y_i = \alpha_i + \mathbf{B}_i \mathbf{y}_1 + \mathbf{\Gamma}_i \mathbf{x}_1 + u_i, \quad (2.27)$$

donde y_i es la i -ésima y de \mathbf{y}_1 , α_i es el correspondiente término de intercepción, \mathbf{B}_i es la fila i -ésima de la matriz \mathbf{B} , $\mathbf{\Gamma}_i$ es la i -ésima fila de la matriz $\mathbf{\Gamma}$, y u_i es el i -ésimo elemento de \mathbf{u} .

Se define \mathbf{A}_i como un vector columna que contiene α_i y todos los elementos distintos de cero de \mathbf{B}_i y $\mathbf{\Gamma}_i$ concatenados en una columna. Sea N el número de casos y \mathbf{Z}_i una matriz de N filas que contiene 1's en la primera columna y las N filas de elementos de \mathbf{y}_1 y \mathbf{x}_1 que tienen coeficientes distintos de cero asociados en las columnas restantes. El vector ($N \times 1$) \mathbf{y}_i contiene los N valores de y_i en la muestra y \mathbf{u}_i es un vector ($N \times 1$) de valores de u_i . De este modo se puede reescribir (2.27) como:

$$\mathbf{y}_i = \mathbf{Z}_i \mathbf{A}_i + \mathbf{u}_i. \quad (2.28)$$

El método de Mínimos Cuadrados Ordinario (OLS) no es apropiado para (2.28) porque en la mayoría de los casos al menos alguna de las variables en \mathbf{Z}_i va a estar correlacionada con \mathbf{u}_i . La única excepción va a ser si \mathbf{Z}_i consiste en x 's perfectamente medidas ($x_j = \zeta_j$) o y 's perfectamente medidas ($y_k = \eta_k$) donde las y 's tampoco están correlacionadas con ζ_i . En una situación diferente a ésta, el método OLS produce estimadores de (2.28) que son no consistentes. Sin embargo, el método de Mínimos Cuadrados en dos etapas (2SLS) provee un estimador consistente de \mathbf{A}_i alternativo.

El 2SLS requiere de variables instrumentales (IV's) para \mathbf{Z}_i . Las IV's deben cumplir las siguientes condiciones:

- (a) estar correlacionadas con \mathbf{Z}_i ,
- (b) no estar correlacionadas con \mathbf{u}_i ,
- (c) ser suficientes en cantidad de modo de disponer de al menos tantas IV's como el número de variables explicativas en el lado derecho de la ecuación.

Generalmente, el conjunto de potenciales IVs está formado por las y 's y las x 's que no están incluidas en \mathbf{Z}_i (excluyendo, por supuesto a y_i). Las excepciones son las variables en \mathbf{Z}_i que no están correlacionadas con u_i , ya que estas variables pueden servir como VIs. Las variables exógenas x 's pueden ser un ejemplo de IVs que pueden aparecer en el lado derecho de la ecuación (2.27).

Para chequear la condición (a) se pueden observar las correlaciones muestrales entre las IVs potenciales y \mathbf{Z}_i . La condición (b) es más difícil de establecer pero la estructura del modelo completo es esencial para evaluarla. Considerando que u_i es igual a $(\epsilon_1 - \mathbf{B}\epsilon_1 - \mathbf{\Gamma}\delta_1 + \zeta_i)$, IVs deben estar no correlacionadas con cada componente de esta variable compuesta. El término $\mathbf{B}\epsilon_1$ excluye a las y 's que dan escala a las variables latentes y que tienen un impacto distinto de cero en y_i . Además, debido a $\mathbf{B}\epsilon_1$, no se puede emplear como IVs a las y 's que tienen errores de medición correlacionados con aquellas y 's en \mathbf{y}_1 que aparecen en la ecuación y_i .

El término $\mathbf{\Gamma}\delta_1$ excluye como variables latentes a las x 's que dan escala a las variables latentes ζ 's y que tienen un impacto distinto de cero en y_i . Asimismo, cualquier x con errores de medida correlacionados con aquellas x 's que aparecen en la ecuación y_i no puede ser IV. Finalmente, el término ζ_i en \mathbf{u}_i elimina cualquiera de las y 's como IV que se correlacione con ζ_i . Esto significa, por ejemplo, que las y 's que son

indicadores de η 's que aparecen luego de η_i en una “cadena causal” no son elegibles para ser IVs.

La condición (c) garantiza que haya suficiente cantidad de IVs para permitir la estimación del modelo. Sin imposición de otras restricciones, esta regla de conteo es una condición necesaria para la identificación de los parámetros en (2.27) y para el empleo de la estimación 2SLS.

Suponiendo que se colectan todas las IVs elegibles para \mathbf{Z}_i y una columna de unos en una matriz \mathbf{V}_i de N filas, la primera etapa de 2SLS es hacer una regresión de \mathbf{Z}_i en \mathbf{V}_i donde (2.29) provee el estimador de los coeficientes:

$$(\mathbf{V}_i' \mathbf{V}_i)^{-1} \mathbf{V}_i' \mathbf{Z}_i, \quad (2.29)$$

Se forma $\hat{\mathbf{Z}}_i$ como:

$$\hat{\mathbf{Z}}_i = \mathbf{V}_i (\mathbf{V}_i' \mathbf{V}_i)^{-1} \mathbf{V}_i' \mathbf{Z}_i, \quad (2.30)$$

La segunda etapa es la regresión por mínimos Cuadrados Ordinarios de y_i en $\hat{\mathbf{Z}}_i$ de modo que

$$\hat{\mathbf{A}} = (\hat{\mathbf{Z}}_i' \hat{\mathbf{Z}}_i)^{-1} \hat{\mathbf{Z}}_i' \mathbf{y}_i, \quad (2.31)$$

La ecuación (2.31) clarifica la importancia y la necesidad de tener suficientes IVs como se describe en condición (c) para las IVs. Si hay menos IVs que variables con coeficientes distintos de cero, entonces $\hat{\mathbf{Z}}_i$ va a ser una matriz de rango incompleto y no va a existir $(\hat{\mathbf{Z}}_i' \hat{\mathbf{Z}}_i)^{-1}$.

Bollen y Bauer (2004) presentan un método para la selección automática de las variables instrumentales, describiendo un algoritmo y su implementación en *softwares* estadísticos de uso común. Los autores enfatizan que las VI implicadas en el modelo están determinadas por el modelo, y que un modelo incorrecto lleva a una selección incorrecta de las VI.

Para emplear el método de estimación de variables instrumentales se debe comenzar por particionar el vector de la variable observada \mathbf{x} en tres partes:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_A \\ \mathbf{x}_B \\ \mathbf{x}_C \end{bmatrix} \quad (2.32)$$

el vector \mathbf{x}_A es un vector ($n \times 1$) y contiene las variables observadas que proveen una escala a las variables latentes, \mathbf{x}_B es una de las x_i 's no incluida en \mathbf{x}_A , y \mathbf{x}_C contiene todo el resto de variables x_i eliminando \mathbf{x}_A y \mathbf{x}_B . El modelo de medida es

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_A \\ \mathbf{x}_B \\ \mathbf{x}_C \end{bmatrix} = \begin{bmatrix} \mathbf{I}_n \\ \boldsymbol{\lambda}_B' \\ \boldsymbol{\Lambda}_C \end{bmatrix} \boldsymbol{\xi} + \boldsymbol{\delta} \quad (2.33)$$

donde \mathbf{I}_n es una matriz identidad ($n \times n$) para las variables observadas que proveen las escalas para ξ , λ'_B es un vector ($1 \times n$) que contiene los coeficientes para variables ξ que influyen en \mathbf{x}_B , y Λ_C es una matriz ($((q - n - 1) \times n)$) de los coeficientes restantes de \mathbf{x}_C . El vector \mathbf{x}_C contiene las variables instrumentales para una ecuación. El estimador IV para λ'_B es

$$\hat{\lambda}'_B = (\mathbf{S}_{AC} \mathbf{S}_{CC}^{-1} \mathbf{S}_{CA})^{-1} \mathbf{S}_{AC} \mathbf{S}_{CC}^{-1} \mathbf{S}_{CB} \quad (2.34)$$

donde \mathbf{S}_{ij} son las matrices de covariancia muestrales de i y j . Las variables que están en \mathbf{x}_B y \mathbf{x}_C rotan dependiendo qué coeficientes son estimados.

Siguiendo el paso anterior se llega a $\hat{\Lambda}_x$. Para los estimadores de $\hat{\Theta}_\delta$ y $\hat{\Phi}$, Häggglund (1982) demuestra que para una matriz $\hat{\Lambda}_x$ dada, el valor de $\hat{\Theta}_\delta$ que minimiza la discrepancia entre la matriz de covariancia muestral \mathbf{S} y la matriz de covariancia implícita $\Sigma(\hat{\theta})$ de acuerdo a los principios de los mínimos cuadrados es

$$\text{diag}(\hat{\Theta}_\delta) = (\mathbf{I} - \mathbf{D} * \mathbf{D})^{-1} \mathbf{g} \quad (2.35)$$

donde \mathbf{D} es igual a $\hat{\Lambda}_x (\hat{\Lambda}'_x \hat{\Lambda}_x)^{-1} \hat{\Lambda}'_x$, “*” significa multiplicación elemento por elemento, y \mathbf{g} es un vector columna ($q \times 1$) formado por los elementos de la diagonal de $(\mathbf{S} - \mathbf{DSD})$. Los elementos de afuera de la diagonal de $\hat{\Theta}_\delta$ son ceros. Usando $\hat{\Theta}_\delta$, de (2.35), Häggglund demuestra que $\hat{\Phi}$ es igual a

$$\hat{\Phi} = (\hat{\Lambda}'_x \hat{\Lambda}_x)^{-1} \hat{\Lambda}'_x (\mathbf{S} - \hat{\Theta}_\delta) \hat{\Lambda}_x (\hat{\Lambda}'_x \hat{\Lambda}_x)^{-1} \quad (2.36)$$

Finalmente, la matriz de covariancia implícita $\Sigma(\hat{\theta})$ es

$$\Sigma(\hat{\theta}) = \hat{\Lambda}_x \hat{\Phi} \hat{\Lambda}'_x + \hat{\Theta}_\delta \quad (2.37)$$

Una vez que se tiene la matriz de covariancias para las variables latentes (η y ξ) se puede aplicar los procedimientos IV o 2SLS, tratando a la matriz de covariancia de la variable latente como si fuera una matriz de covariancia de variables observadas. Esto lleva a estimar \mathbf{B} y Γ . Un estimador de Φ es parte de la matriz de covariancia de η y ξ .

Finalmente, un estimador de Ψ es

$$\hat{\Psi} = (\mathbf{I} - \hat{\mathbf{B}})' \hat{\Sigma}_{\eta\eta} (\mathbf{I} - \hat{\mathbf{B}}) - \hat{\Gamma} \hat{\Phi} \hat{\Gamma}' \quad (2.38)$$

2.7.2.2 Pasos en la secuencia

Una vez que se tienen los valores iniciales de los parámetros, se debe establecer el criterio para moverse de un valor al siguiente, esto es, de $\theta^{(i)}$ a $\theta^{(i+1)}$. El criterio básico es que, en la medida que se avance a través de la secuencia $\theta^{(1)}$, $\theta^{(2)}$, ..., $\theta^{(f)}$ los valores de la función de ajuste deben declinar, con la función de ajuste en un mínimo para $\theta^{(f)}$.

Idealmente, en cada paso $F(\hat{\theta}^{(i+1)})$ es menor que $F(\hat{\theta}^{(i)})$, donde $F(\cdot)$ representa el valor de la función de ajuste en los valores dados en los pasos. Sin embargo, aunque las funciones de ajuste sean buenas, no siempre decrecen monotónicamente.

El gradiente de la función, $\frac{\partial F_{ML}}{\partial \hat{\theta}}$, juega un rol en la mayoría de los procedimientos de minimización. Para un vector $\hat{\theta}$ de estimadores de parámetros desconocidos, el valor $\hat{\theta}^{(i+1)}$ para muchos de estos procedimientos es

$$\hat{\theta}^{(i+1)} = \hat{\theta}^{(i)} - \mathbf{C}^{(i)} \mathbf{g}^{(i)} \quad (2.39)$$

donde $\mathbf{g}^{(i)}$ es el vector gradiente $\frac{\partial F_{ML}}{\partial \hat{\theta}}$ en $\hat{\theta}^{(i)}$, y $\mathbf{C}^{(i)}$ es una matriz definida positiva.

Muchos de los métodos de minimización se distinguen por su elección de la matriz $\mathbf{C}^{(i)}$. La elección más simple es una matriz identidad. Esto lleva al método del descenso más abrupto (*steepest descent*). Se puede ajustar el “largo del paso” multiplicando $\mathbf{C}^{(i)}$ por una constante (por ej. 1/2) para alterar el cambio resultante en $\hat{\theta}^{(i+1)}$. Las mayores desventajas de este método son que es muy lento y no es muy sensible a los cambios de forma que F_{ML} puede tomar.

Otra elección para $\mathbf{C}^{(i)}$ es la inversa de las derivadas parciales segundas, esto es la inversa de la matriz hessiana, de F_{ML} con respecto a θ o $\left[\frac{\partial^2 F_{ML}}{\partial \hat{\theta} \partial \hat{\theta}} \right]^{-1}$ en $\hat{\theta}^{(i)}$. Esta elección está basada en una expansión de la serie de Taylor de F_{ML} de $\hat{\theta}$ alrededor de $\hat{\theta}^{(i)}$. La principal desventaja de este algoritmo de Newton-Raphson es que se requiere la primera y segunda derivadas de F_{ML} con respecto a $\hat{\theta}$. Asimismo, los requerimientos computacionales son grandes ya que la inversa de las derivadas parciales segundas es calculada en cada paso. En general, $\hat{\theta}^{(i+1)}$ es determinada por

$$\hat{\theta}^{(i+1)} = \hat{\theta}^{(i)} - \left[\frac{\partial^2 F_{ML}}{\partial \hat{\theta} \partial \hat{\theta}} \right]^{-1} \left[\frac{\partial F_{ML}}{\partial \hat{\theta}} \right] \quad (2.40)$$

SAS, en sus procedimientos CALIS y TCALIS emplea una variedad de algoritmos no-lineales de optimización para estimar los parámetros entre los que se encuentran los dos mencionados arriba. El criterio de calcular el gradiente lleva mayor tiempo de procesamiento que emplear el valor de la función, y calcular la matriz hessiana mucho más tiempo y memoria que el método del gradiente, especialmente cuando hay muchos parámetros para estimar. Lamentablemente, las técnicas de optimización que no emplean la matriz hessiana requieren más iteraciones y son generalmente más lentas. Asimismo, las técnicas que no usan la matriz hessiana son menos confiables porque pueden terminar en un óptimo local en vez del óptimo global.

2.7.2.3 Criterio de finalización

Existen varios criterios de finalización de las iteraciones. Por ejemplo, si la diferencia en la función de ajuste de una iteración a la siguiente es menor que algún valor muy pequeño, se pueden parar las iteraciones. O si hay muy poca diferencia entre las estimaciones de los parámetros de una iteración a la siguiente, se dice que el método converge. Se pueden, también establecer límites en el número de iteraciones o en el tiempo de procesamiento.

2.8. Errores estándar aproximados

Con excepción de la estimación por mínimos cuadrados no-ponderados, los errores estándar aproximados pueden ser calculados como los elementos de la diagonal de la matriz

$$\frac{2}{N-1} \mathbf{H}^{-1}, \quad (2.41)$$

donde la matriz \mathbf{H} es la matriz hessiana aproximada de F evaluada en las estimaciones finales, y N es el tamaño de la muestra.

Si la matriz de correlación o covariancia dada es singular, el software SAS ofrece dos formas de calcular la inversa generalizada de la matriz de información, y, por lo tanto dos formas de calcular los errores estándar de los estimadores implícitamente restringidos, los valores del estadístico t y los índices de modificación. Una forma es la inversa de Moore-Penrose y otra es la inversa G2. La primera forma calcula un estimador del espacio nulo usando una descomposición de los autovalores. La segunda forma es producida barriendo las filas y columnas linealmente independientes y llevando a cero las dependientes.

2.9 Bondad de Ajuste del Modelo

Los coeficientes estimados y la fuerza de las asociaciones requieren ser cuidadosamente examinados. Las medidas para evaluar el ajuste del modelo pueden ser agrupadas en medidas globales y medidas individuales.

La evaluación de la bondad del ajuste global para los modelos de ecuaciones estructurales no es tan directa como en otras técnicas. No se ha desarrollado en SEM una única prueba que sea la que mejor describa la fuerza de las predicciones del modelo. En su lugar, los investigadores han desarrollado varias medidas de bondad de ajuste, que, cuando se usan en combinación, evalúan el resultado desde tres perspectivas: ajuste conjunto, ajuste comparativo respecto a un modelo base y la parsimonia del modelo (Hair *et al.*, 1999).

Ninguna de las medidas que se presentan en este trabajo, con excepción del estadístico de χ^2 , tiene un contraste estadístico asociado. Aunque se han sugerido algunas líneas maestras, no existe una prueba absoluta, y el investigador debe decidir en última instancia si el nivel de ajuste alcanzado por el modelo es aceptable.

Como se mencionó al comienzo del capítulo, la hipótesis de la estructura de covariancia es que $\Sigma = \Sigma(\theta)$. Las medidas de ajuste globales ayudan a evaluar si esta hipótesis es válida y si no, a medir la desviación entre Σ y $\Sigma(\theta)$. Como Σ y $\Sigma(\theta)$ son parámetros poblacionales desconocidos, se examinan sus contrapartes muestrales, S y $\Sigma(\hat{\theta})$. La matriz S es la matriz de covariancia muestral usual, y $\Sigma(\hat{\theta})$ es la matriz de covariancia implicada evaluada en la estimación de θ que minimiza, ya sea F_{ML} , F_{GLS} o F_{ULS} , y que puede ser abreviada como $\hat{\Sigma}$.

Prácticamente, todas las medidas de ajuste global involucran S y $\hat{\Sigma}$. Estos índices evalúan qué tan cerca está S de $\hat{\Sigma}$ de diferentes maneras. La principal ventaja de las medidas de ajuste global es que evalúan el modelo total y que pueden indicar fallas que no son reveladas por el ajuste de los componentes del modelo (ecuaciones y estimadores de parámetros). Una de las limitantes es que no pueden ser aplicados a modelos exactamente identificados ya que, en esta situación, S siempre es igual a $\hat{\Sigma}$. En modelos sobreidentificados es posible que S difiera de $\hat{\Sigma}$ por lo que se puede probar si las restricciones que llevan a la sobreidentificación son válidas. Una segunda limitante de las medidas de ajuste globales es que pueden diferir del ajuste de los componentes del modelo. Por ejemplo, el ajuste global puede ser bueno, pero los estimadores de los parámetros pueden no ser estadísticamente significativos o pueden tener signos opuestos a lo esperado. Asimismo, el ajuste global de $\hat{\Sigma}$ y S no indica qué tan bien las variables explicativas predicen a las variables dependientes. Por estas razones no se pueden emplear las medidas de ajuste global aisladas de las medidas de ajuste de los componentes del modelo.

2.9.1 Medidas de ajuste global

Las medidas de bondad de ajuste global evalúan el grado en que el modelo conjunto (modelo estructural y de medida) predice la matriz de correlación o covariancia observada. No se hace distinción acerca de si el modelo de ajuste es mejor o peor en los modelos de medida o estructural.

2.9.1.1 Residuales

Cuando la hipótesis de la estructura de covariancia es cierta, la matriz poblacional de covariancia de los residuales $\Sigma - \Sigma(\theta)$ es una matriz de ceros. Cualquier residual en la población que sea distinto de cero está indicando que la especificación del modelo está en error. Si bien no se cuenta directamente con las matrices poblacionales, se tiene la contraparte muestral $S - \hat{\Sigma}$.

La matriz de residuales es la forma más sencilla para evaluar el ajuste global. Las covariancias de residuales individuales son $(s_{ij} - \hat{\sigma}_{ij})$, donde s_{ij} es el elemento ij -ésimo en S y $\hat{\sigma}_{ij}$ es el correspondiente en $\hat{\Sigma}$. Un residual positivo indica que el modelo subestima la covariancia entre dos variables, mientras que uno negativo indica que la covariancia predicha es muy alta. Los residuales individuales, así como la media y mediana de sus valores absolutos pueden ayudar a evaluar el ajuste del modelo.

La situación ideal es que todos los residuales estén cercanos a cero en un “buen” modelo. Pero los residuales en la muestra pueden ser afectados por varios factores: (1) la diferencia entre Σ y $\Sigma(\theta)$, (2) las escalas de las variables observadas, y (3) las fluctuaciones del muestreo. El interés principal es emplear los residuales para detectar (1). Cuando ocurre que $\Sigma \neq \Sigma(\theta)$, una o más de las covariancias o variancias de las variables observadas no está siendo exactamente predicha por el modelo.

La segunda causa se refiere a que las magnitudes de los residuales individuales y su media son alterados si las variables observadas están medidas en diferentes unidades. Un residual grande puede ser debido a una variable observada con unidades de escala que tienen mucho mayor rango que el de otras variables.

Una solución simple es calcular los residuales de las correlaciones:

$$r_{ij} - \hat{r}_{ij} \quad (2.42)$$

donde r_{ij} es la correlación muestral entre las variables i -ésima y j -ésima, y \hat{r}_{ij} es la predicha por el modelo correspondiente. Empleando la definición de correlación, \hat{r}_{ij} es $\frac{\hat{\sigma}_{ij}}{\sqrt{\hat{\sigma}_{ii}\hat{\sigma}_{jj}}}$, donde $\hat{\sigma}_{ij}$ es el elemento ij -ésimo de $\hat{\Sigma}$ y $\hat{\sigma}_{ii}$ y $\hat{\sigma}_{jj}$ son las variancias de las

variables i -ésima y j -ésima, respectivamente, de la matriz. Como r_{ij} y \hat{r}_{ij} van de -1 a +1, los residuales de correlaciones tienen un rango teórico de -2 a +2. Un valor cercano a estos límites indica serios problemas de ajuste. Un residual de correlación debería estar muy cerca de cero para que el modelo se considere bien ajustado. A su vez, la media y la mediana de los valores absolutos de los residuos también permiten evaluar la magnitud general de los residuales.

El tercer factor que afecta a los residuales de la muestra y los residuales de las correlaciones, es el error de muestreo. Aún cuando $\Sigma = \Sigma(\theta)$ sea cierta, la magnitud esperada de los residuales de la muestra dependen de N . Bajo condiciones generales adecuadas S converge a Σ y $\hat{\Sigma}$ converge a $\Sigma(\theta)$ en la medida en que N se va haciendo mayor. Similarmente, $S - \hat{\Sigma}$ converge a $\Sigma - \Sigma(\theta)$ cuando $N \rightarrow \infty$. Para un modelo dado, $(s_{ij} - \hat{\sigma}_{ij})$ tiende a ser menor cuanto mayor es la muestra (Browne, 1982). La implicancia de esto es que, al juzgar residuales en muestras más pequeñas, se esperan residuales mayores (en el promedio) que cuando se examinan residuales en muestras grandes, siendo el modelo verdadero en ambos casos.

Jöreskog y Sörbom (1986) proponen un *residuo normalizado* que provee una corrección aproximada por estos efectos de tamaño de muestra y diferencias de escala. La expresión es:

$$\text{residuo normalizado} = \frac{(s_{ij} - \hat{\sigma}_{ij})}{\left[(\hat{\sigma}_{ii}\hat{\sigma}_{jj} + \hat{\sigma}_{ij}^2) / N \right]^{1/2}} \quad (2.43)$$

El numerador es el residual y el denominador es la raíz cuadrada de su variancia asintótica estimada. Los mayores valores absolutos de los residuales normalizados indican los elementos s_{ij} peor ajustados del modelo. Jöreskog y Sörbom (1988) señalaron que la variancia asintótica estimada en el denominador tiende a ser muy alta. Esto implica que los residuales normalizados son menores que una variable normal estándar.

2.9.1.2 Prueba de χ^2

Browne (1974, 1982, 1984) provee una justificación para emplear $(N - 1)F_{\text{GLS}}$ o $(N - 1)F_{\text{ML}}$ como estimadores χ^2 basado en principios de mínimos cuadrados generalizados. La hipótesis nula, $H_0: \Sigma = \Sigma(\theta)$, significa que la especificación de los parámetros fijos, libres y restringidos en Λ_x , Φ , Θ_δ es válida. Bajo H_0 se estiman por ML los parámetros libres y restringidos en esas matrices los que, junto con los parámetros fijos, comprenden las matrices $\hat{\Lambda}_x$, $\hat{\Phi}$, y $\hat{\Theta}_\delta$. Luego, se forma $\hat{\Sigma} (= \Sigma(\hat{\theta}))$, que es la matriz de covariancia predicha de las variables observadas en la muestra bajo H_0 . El logaritmo de la función de verosimilitud, $\ln L_0$, corresponde a esta hipótesis nula. Cuando se evalúa en \mathbf{S} y $\hat{\Sigma}$, el $\ln L_0$ es

$$\ln L_0 = -\frac{N-1}{2} \left\{ \ln |\hat{\Sigma}| + \text{tr}(\hat{\Sigma}^{-1}\mathbf{S}) \right\} \quad (2.44)$$

La expresión (2.44) es el numerador de la prueba del cociente de verosimilitud. Para formar el cociente de verosimilitudes se necesita una hipótesis alternativa, H_1 . Se busca una H_1 en la que la función de verosimilitud, L_1 , esté en un máximo. Una H_1 menos restrictiva puede ser que $\hat{\Sigma}$ sea cualquier matriz definida positiva. Si se emplea \mathbf{S} como $\hat{\Sigma}$, el $\ln L_1$ está en su valor máximo. Ignorando las constantes irrelevantes de la función de verosimilitud para H_1 , el $\ln L_1$, es

$$\begin{aligned} \ln L_1 &= -\frac{N-1}{2} \left\{ \ln |\mathbf{S}| + \text{tr}(\mathbf{S}^{-1}\mathbf{S}) \right\} \\ &= -\frac{N-1}{2} \left\{ \ln |\mathbf{S}| + (p+q) \right\} \end{aligned} \quad (2.45)$$

La expresión (2.45) es el logaritmo natural del denominador del cociente de verosimilitud.

La hipótesis H_1 no es informativa en sí misma, ya que no representa ningún modelo en el sentido que lo hace H_0 . La elección de H_1 de modo que $\hat{\Sigma} = \mathbf{S}$ provee un estándar de ajuste perfecto contra el que comparar H_0 .

El logaritmo natural del cociente de verosimilitud, $\ln(L_0/L_1)$, multiplicado por -2 se distribuye como una variable χ^2 cuando H_0 es verdadera y $(N - 1)$ es grande. En este caso

$$\begin{aligned} -2 \ln \left(\frac{L_0}{L_1} \right) &= -2 \ln L_0 + 2 \ln L_1 \\ &= (N-1) \left[\ln |\hat{\Sigma}| + \text{tr}(\hat{\Sigma}^{-1}\mathbf{S}) \right] - (N-1) \left[\ln |\mathbf{S}| + (p+q) \right] \\ &= (N-1) \left(\ln |\hat{\Sigma}| + \text{tr}(\hat{\Sigma}^{-1}\mathbf{S}) - \ln |\mathbf{S}| - (p+q) \right) \end{aligned} \quad (2.46)$$

Como puede observarse la expresión entre paréntesis de (2.46) es la función de ajuste F_{ML} de (2.15) evaluada en S y $\hat{\Sigma}$. La expresión (2.46) muestra que $(N - 1)$ veces la función de ajuste F_{ML} evaluada en $\hat{\theta}$ está aproximadamente distribuida como una variable χ^2 . Los grados de libertad son $(1/2)(p + q)(p + q + 1) - t$, donde el primer término es el número de elementos no redundantes en S dadas $p + q$ variables observadas, y t es el número de parámetros libres en θ .

A diferencia de las pruebas de hipótesis en regresión, para la prueba de χ^2 , en este caso, la hipótesis nula H_0 es que las restricciones en Σ dadas por el modelo son válidas. Por esta razón, cuanto mayor sea el valor p de la prueba mayor será el ajuste del modelo.

Existen algunas consideraciones que se deben tener en cuenta en el empleo de la prueba de χ^2 . La aproximación de χ^2 supone que: (1) las variables observadas tienen distribución Normal multivariada, (2) se analiza la matriz de covariancia, (3) la muestra es suficientemente grande, y (4) la hipótesis $H_0: \Sigma = \Sigma(\theta)$ se cumple en forma perfecta. En la práctica, uno o más de estos supuestos son violados. Considerando el primer supuesto, Browne (1974, 1982) muestra que $(N - 1)F_{ML}$ y $(N - 1)F_{GLS}$ son estimadores χ^2 cuando las variables observadas no tienen excesiva curtosis. En muchas situaciones prácticas es común encontrar variables no normales con excesiva curtosis. Browne (1984) sugiere que distribuciones leptocúrticas llevan a muchos rechazos incorrectos de la hipótesis nula. Por otro lado, si los datos tienen una distribución platicúrtica, el análisis puede fallar en rechazar modelos de ajuste pobre (Anderson y Gerbing, 1988).

El segundo factor que puede influenciar la prueba de χ^2 es si se emplea la matriz de covariancia o la matriz de correlación en el análisis. Existe cierta discusión sobre este punto. Por un lado, Jöreskog y Sörbom (1986) manifiestan que el estimador χ^2 no es fidedigno cuando se analiza la matriz de correlación. Por otro lado, basado en un trabajo de simulación, Boomsma (1983) establece que bajo condiciones de invariancia estandarizando las variables observadas a una variancia de 1 no afecta a la estimación χ^2 para la bondad de ajuste. La "condición de invariancia" es violada cuando, por ejemplo, las cargas de los factores o variancias del error son forzados a ser iguales.

Con respecto al tercer factor, McQuitty (2004) destacó la importancia de determinar el tamaño mínimo de la muestra de modo de obtener una potencia estadística adecuada para un modelo dado, previo a la colección de datos. Schreiber *et al.* (2006) mencionaron que, a pesar de que el tamaño de muestra necesario depende de la normalidad de los datos y del método de estimación que se emplee, el tamaño de muestra de aceptación general es de 10 veces el número de parámetros libres estimados en el modelo. Lamentablemente, en los modelos de ecuaciones estructurales una muestra muy grande resulta en un estadístico χ^2 con tal potencia que puede demostrar significancia estadística cuando existen solo diferencias triviales entre las covariancias predichas y observadas. De este modo, la prueba de χ^2 puede resultar en el rechazo de modelos que parecen ajustar bastante bien el modelo (Mulaik *et al.*, 1989). Por otra parte, Boomsma (1983) en su trabajo de simulación sugiere que el estimador χ^2 . $(N - 1)F_{ML}$ no es fidedigno para muestras menores a 50 y recomienda 100 o más casos. También en un trabajo de simulación, Anderson y Gerbing (1984) encuentran grandes desviaciones entre $(N - 1)F_{ML}$ y la variable χ^2 en muestras pequeñas (< 100). En ambos estudios, esta desviación de una χ^2 resulta frecuentemente en demasiados rechazos de la hipótesis nula.

Finalmente, el cuarto factor que afecta la aproximación χ^2 es que se asume que $\Sigma = \Sigma(\theta)$ es exactamente cierta. En general, no se espera tener una descripción exacta de la realidad. El objetivo es más modesto, en el sentido de que si el modelo que lleva a $\Sigma(\theta)$

ayuda a entender la relación entre las variables y logra una tarea razonable en la correspondencia con los datos, se puede considerar como parcialmente validado. El supuesto de que se ha identificado al proceso exacto que genera los datos no puede ser aceptado. Sin embargo, la prueba de χ^2 se deriva de una comparación entre un modelo hipotético H_0 con H_1 , que es un modelo de ajuste perfecto. Un ajuste perfecto puede ser un estándar inapropiado, y un estimador χ^2 alto puede indicar lo que ya se sabe, que H_0 se cumple aproximadamente pero no en forma perfecta.

En síntesis, la prueba de χ^2 tiene ventajas y desventajas como medida resumen del ajuste del modelo. Cuando el tamaño de la muestra es adecuadamente grande, se analiza la matriz de covariancia, las variables observadas no tienen excesiva curtosis y H_0 es verdadera, entonces

$(N - 1)F_{ML}$ es una buena aproximación a una variable χ^2 adecuada para pruebas de significancia estadística. Sin embargo, si una o más de estas condiciones son violadas, entonces la prueba de χ^2 pierde algo de su valor. Por esta razón, se recomienda que se trate al estadístico χ^2 como un índice de bondad de ajuste general pero no como una prueba estadística en el sentido estricto (Jöreskog y Sörbom, 1988). En este sentido, es conveniente suplementar la prueba de χ^2 con una serie de otros índices de bondad de ajuste. En algunos casos, estos índices pueden revelar un ajuste relativamente bueno aún cuando la prueba de χ^2 haya resultado en el rechazo del modelo.

Como se mencionó, la prueba usual de ajuste exacto es invariablemente falsa en situaciones prácticas y seguramente va a ser rechazada si la muestra es suficientemente grande. En este sentido, Steiger (2004) reflexiona acerca de la conveniencia de evaluar qué tan bueno (o qué tan malo) es el modelo con un intervalo de confianza, en lugar de probar que el modelo ajusta en forma perfecta. Steiger y Lind (1980) desarrollaron el error de aproximación cuadrático medio (RMSEA). Este índice es una medida de ajuste aproximado:

$$RMSEA = \sqrt{\max\left(\frac{F}{df} - \frac{1}{N-1}, 0\right)} \quad (2.47)$$

Los límites inferior y superior del intervalo de confianza son calculados empleando la distribución acumulada de la χ^2 no-centrada $\Phi(x/\lambda, df) = \alpha$, con $x = (N - 1)F$, λ_L satisfaciendo

$\Phi(x/\lambda_L, df) = 1 - (\alpha/2)$, y λ_U satisfaciendo $\Phi(x/\lambda_U, df) = (\alpha/2)$:

$$(RMSEA_L; RMSEA_U) = \left(\sqrt{\frac{\lambda_L}{(N-1)df}}; \sqrt{\frac{\lambda_U}{(N-1)df}} \right) \quad (2.48)$$

2.9.1.3 Índices de ajuste incremental

Uno de los índices de ajuste más populares es el de ajuste normado (NFI) propuesto por Bentler y Bonett (1980) como alternativa a la prueba χ^2 :

$$NFI = \frac{F_b - F_m}{F_b}$$

$$= \frac{\chi_b^2 - \chi_m^2}{\chi_b^2} \quad (2.49)$$

En (2.49) F_b es el valor de la función de ajuste de un “modelo base” y F_m es el valor de la función de ajuste para el modelo de la hipótesis (F puede ser F_{ML} , F_{GLS} o F_{ULS}). La fórmula alternativa con χ^2 es, a veces, computacionalmente más conveniente. Es equivalente a la otra fórmula, con χ_b^2 siendo el estadístico χ^2 para el modelo base y χ_m^2 el correspondiente al modelo propuesto.

El modelo base puede ser el nulo, en el que solo se consideran las variancias de las variables observadas. Los valores del NFI van de 0 a 1, con valores mayores a 0.90 indicando un ajuste aceptable del modelo a los datos. Este índice puede ser interpretado como la proporción de la covariación observada que es explicada por un modelo de medida o estructural dado (comparado con un modelo general nulo), o como la mejora “incremental” en el ajuste del modelo propuesto en relación a uno base.

Una limitante del NFI es que no tiene en cuenta los grados de libertad. El valor de F_m puede ser reducido agregando parámetros, en forma análoga al incremento del R^2 en la regresión cuando se agregan parámetros al modelo. El χ_m^2 (o F_m) con muchos parámetros estimados y mucha complejidad nunca es mayor y casi siempre es menor que el χ_m^2 correspondiente a modelos más parsimoniosos que contienen subconjuntos de estos parámetros. El modelo complejo va a tener un NFI mayor aún con menos grados de libertad.

Bollen (1988) citado por Bollen (1989) propone una modificación de NFI que disminuye la dependencia de N de la media de la distribución de este índice y que tiene en cuenta los grados de libertad, df_m . Se lo conoce como Índice de Ajuste No Normado Delta2 (NNFI):

$$\begin{aligned} NNFI &= \frac{F_b - F_m}{F_b [df_m / (N - 1)]} \\ &= \frac{\chi_b^2 - \chi_m^2}{\chi_b^2 - df_m} \end{aligned} \quad (2.50)$$

El razonamiento en que se basa (2.50) es que el valor esperado de χ_m^2 es el estándar de mejor ajuste. Si la distribución del estadístico para el modelo probado es la χ^2 aproximada, entonces su esperanza es igual a df_m . Para un modelo correcto, el numerador ($\chi_b^2 - \chi_m^2$) es en promedio ($\bar{\chi}_b^2 - df_m$), donde $\bar{\chi}_b^2$ es el promedio del χ^2 correspondiente al modelo base. En general $\bar{\chi}_b^2$ no sigue una distribución χ^2 ya que se sabe que el modelo base es demasiado restrictivo. Bajo ciertas condiciones puede seguir una distribución χ^2 no-centrada. Si lo que se espera para un modelo correcto es, en promedio, ($\bar{\chi}_b^2 - df_m$) esta expresión provee un estándar alternativo con el que comparar ($\chi_b^2 - \chi_m^2$). Estas dos expresiones deberían ser iguales en el promedio en un modelo correcto. El χ_b^2 reemplaza a $\bar{\chi}_b^2$ en el denominador de (2.50) porque es el único estimador disponible de $\bar{\chi}_b^2$.

Otra medida de ajuste es la propuesta por Bollen (1986), conocido como Índice Normado Rho1:

$$\rho_1 = \frac{\left(\frac{F_b}{df_b}\right) - \left(\frac{F_m}{df_m}\right)}{\left(\frac{F_b}{df_b}\right)} = \frac{\left(\frac{\chi_b^2}{df_b}\right) - \left(\frac{\chi_m^2}{df_m}\right)}{\left(\frac{\chi_b^2}{df_b}\right)} \quad (2.51)$$

Los modelos que tienen menores funciones de ajuste con, relativamente, pocos parámetros tienen mayor ρ_1 que modelos con el mismo ajuste pero especificaciones más complejas. Como el valor más bajo que puede tomar F_m es 0, el valor mayor que puede tomar ρ_1 es 1.

Bentler (1989) propuso un índice de ajuste conocido como Índice de Ajuste Comparativo (CFI) que compara los valores de los χ^2 del modelo propuesto y del modelo bajo independencia considerando los grados de libertad de ambos modelos:

$$CFI = 1 - \frac{\max(\chi_m^2 - df_m, 0)}{\max(\chi_0^2 - df_0, 0)} \quad (2.52)$$

donde el subíndice 0 indica el modelo bajo independencia.

La selección del modelo de base afecta las medidas de ajuste presentadas arriba. Cuanto más restrictivo es el modelo de base, mejor es el ajuste del modelo propuesto con relación al de base. Esto tiende a inflar los índices incrementales de ajuste comparados con los correspondientes a un modelo de base menos restrictivo. Una de las formas de disminuir este problema es siempre reportar el ajuste del modelo con el que no considera factores.

2.9.1.4 Otras medidas de ajuste

Jöreskog y Sörbom (1986) proponen un índice de bondad de ajuste (GFI), considerando, también un GFI ajustado por los grados de libertad para modelos ajustados por F_{ML} :

$$GFI_{ML} = 1 - \frac{tr\left[\left(\hat{\Sigma}^{-1}\mathbf{S} - \mathbf{I}\right)^2\right]}{tr\left[\left(\hat{\Sigma}^{-1}\mathbf{S}\right)^2\right]} \quad (2.52)$$

$$AGFI_{ML} = 1 - \left[\frac{(p+q)(p+q+1)}{2df}\right] \cdot [1 - GFI_{ML}] \quad (2.53)$$

El GFI_{ML} mide la cantidad relativa de las variancias y covariancias en \mathbf{S} que son predichas por $\hat{\Sigma}$. El $AGFI_{ML}$ ajusta por los grados de libertad de un modelo relativo al número de variables observadas. Para un valor dado de GFI_{ML} y $(p+q)$, $AGFI_{ML}$

beneficia a los modelos más simples, con menos parámetros. Ambos índices llegan a un máximo de 1 cuando $\mathbf{S} = \hat{\Sigma}$.

Mulaik *et al.* (1989) proponen el índice de bondad de ajuste de parsimonia (PGFI) como una modificación del GFI que tiene en cuenta la parsimonia del modelo:

$$PGFI = \frac{df_{\min}}{df_0} GFI \quad (2.54)$$

donde df_{\min} son los grados de libertad del modelo propuesto y df_0 son los correspondientes al modelo bajo independencia.

El PGFI emplea el mismo factor de parsimonia que el índice de ajuste de parsimonia normado de Bentler- Bonett (James *et al.*, 1982). La parsimonia se define en la consecución de mayores niveles de ajuste por grado de libertad utilizado (un grado de libertad por parámetro estimado), por lo que es deseable una mayor parsimonia. Valores elevados de PGFI son deseables y su uso principal consiste en la comparación de modelos con diferentes grados de libertad. Si bien no existen niveles recomendados de ajuste, cuando se comparan modelos, se proponen diferencias de 0.06 a 0.09 como indicativas de diferencias sustanciales entre los modelos (Williams y Holahan, 1994)

Jöreskog y Sörbom (1986) proponen índices similares para modelos ajustados con F_{ULS} , y Tanaka y Huba (1985) para los ajustados con F_{GLS} .

Otro índice propuesto es el N Crítico (CN) de Hoelter (1983):

$$CN = \frac{\chi^2_{\text{crítico}}}{F} + 1 \quad (2.55)$$

donde el χ^2 crítico es el valor crítico de la distribución χ^2 con los grados de libertad correspondientes al modelo probado y un nivel de significación dado. F es el valor de F_{ML} o F_{GLS} , en \mathbf{S} y $\hat{\Sigma}$. CN da el tamaño de muestra para el cual el valor de F llevaría al rechazo de H_0 con un dado nivel de significación. Hoelter (1983) sugiere como punto de corte el valor $CN \geq 200$. Como N no entra en la fórmula de CN, su valor es el mismo para un χ^2 crítico y valor F dados para todos los tamaños de muestra. Sin embargo, cuando H_0 es válida, F tiende a 0 cuando N aumenta y la media de la distribución muestral de CN aumenta con N . Bollen y Liang (1988) demuestran que puede ocurrir una asociación positiva también en algunos modelos mal especificados. Así, CN puede llevar a una evaluación del ajuste excesivamente pesimista en caso de muestras pequeñas.

Una medida de ajuste frecuentemente empleada es el estadístico χ^2 dividido sus grados de libertad. Su justificación está dada porque el valor esperado de la variable χ^2 es igual a sus grados de libertad. De modo que χ^2/df estima cuántas veces su valor esperado es el estimador χ^2 cuando $(N - 1)F_{\text{ML}}$ o $(N - 1)F_{\text{GLS}}$ se aproximan a una distribución χ^2 . No existe un consenso de cuál es el valor para un buen ajuste, llegando a valores de 2 o 3 o menos (Carmines y McIver, 1981) hasta valores de hasta 5.

Tres índices que comparan el ajuste de modelos son: el índice de Akaike (1974, 1987), el criterio consistente de Akaike (Bozdogan, 1987) y el criterio Bayesiano de Schwarz (Schwarz, 1978; Sclove, 1987). El índice de Akaike es:

$$AIC = \chi_m^2 - 2df \quad (2.56)$$

donde df es el número de grados de libertad y χ_m^2 es el valor del estadístico χ^2 para el modelo probado.

El criterio propuesto por Bozdogan (1987) es preferido por algunos investigadores al de Akaike:

$$CAIC = \chi_m^2 - df (\ln(N) + 1) \quad (2.57)$$

El índice de Schwarz es, también, una modificación del criterio de Akaike y es:

$$SBC = \chi_m^2 - df \ln(N) \quad (2.58)$$

La recomendación es elegir aquel modelo con el menor valor de cada uno de estos criterios. Los tres índices penalizan los modelos con muchos parámetros, y CAIC y SBC, además, ponderan los grados de libertad por el logaritmo del tamaño de muestra, por esta razón son considerados medidas de ajuste de parsimonia.

Otro índice cercano al AIC es el Índice Esperado de Validación Cruzada (ECVI) de Browne y Cudek (1989, 1993). El ECVI es realmente un parámetro que se estima por el estadístico:

$$c = F \left[\mathbf{S}, \Sigma(\hat{\boldsymbol{\theta}}) \right] + \frac{2t}{N-1} \quad (2.59)$$

donde $F(\mathbf{S}, \Sigma(\hat{\boldsymbol{\theta}}))$ es el mínimo valor de la función de ajuste, t es el número de parámetros y N es el tamaño de muestra.

Mientras que AIC fue derivado de la teoría de información, ECVI es una medida de la discrepancia entre la matriz de covariancia implicada de acuerdo al modelo en la muestra analizada y la matriz de covariancia que se esperaría en otra muestra del mismo tamaño. De este modo, ECVI evalúa la bondad de ajuste que conseguiría el modelo estimado en otra muestra del mismo tamaño (Kaplan, 2000). Cuando se comparan varios modelos, el menor ECVI estimado indica el modelo de mejor ajuste.

2.9.2 Medidas de Ajuste de los componentes del modelo

Como se mencionó anteriormente, se pueden obtener resultados sin sentido para los parámetros individuales junto con buenas medidas de ajuste global. Asimismo, las medidas resumen no son siempre aplicables. Para modelos exactamente identificados, la matriz $\hat{\Sigma}$ es siempre igual a \mathbf{S} , por lo que el χ^2 es cero. Por estas razones es fundamental examinar los componentes del modelo.

2.9.2.1 Estimadores de los parámetros

Generalmente, el signo y la magnitud de los parámetros desconocidos en Λ_x , Λ_y , ψ , Φ , Θ_δ , Θ_ϵ son de interés. El primer paso en el examen de los estimadores es observar que éstos tengan sentido. Los coeficientes estandarizados pueden ser útiles en la comparación de efectos cuando las variables observadas o latentes tienen diferentes unidades de medida. Si las estimaciones de los parámetros, estandarizadas o no

estandarizadas se alejan mucho de lo esperado, la hipótesis puede ser defectuosa o el modelo puede estar mal especificado. Otro de los signos de una mala especificación del modelo es la aparición de una solución impropia. Las soluciones impropias se refieren a estimaciones muestrales que toman valores imposibles en la población. Un caso común, conocido como “caso de Heywood”, es cuando las variancias son negativas. Otra situación que puede ocurrir es cuando la correlación entre dos variables es mayor que 1.

Las soluciones impropias pueden ser causadas por diversos factores. El parámetro poblacional puede ser un valor aceptable pero que está cercano al límite de los valores admisibles. En estas situaciones la estimación puede tomar un valor inadmisble debido a fluctuaciones del muestreo. Una forma de controlar esto es verificar que esos valores límite son razonables para esos parámetros. Por ejemplo, preguntarse si es posible que las variables involucradas tengan una correlación perfecta o que la variancia del error sea cero. Si estas preguntas no pueden contestarse afirmativamente, pueden existir otras explicaciones para que ocurran este tipo de resultados.

En primer lugar, la muestra pudo haber sido una muestra atípica poco probable. En segundo lugar, la matriz de covariancia (correlación) analizada pudo haber tenido valores atípicos (*outliers*) u observaciones influyentes que llevaron a medidas de asociación distorsionadas de las variables observadas. En estos casos se pueden seguir distintas estrategias. Una es re-estimar el modelo con la restricción de que ninguna de las variancias del error sea negativa. Otra estrategia es eliminar la variable con variancia negativa. Finalmente, otra posible acción es ignorar el valor negativo y considerarlo esencialmente 0. El problema de eliminar una de las variables observadas es que puede llevar a un modelo no identificado a menos que se agreguen restricciones adicionales. Sin duda, uno de los primeros controles que se deben realizar en estos casos es verificar que no existan valores atípicos o influyentes que, al eliminarlos, solucionan el problema.

En trabajos de simulación, Boomsna (1982) y Anderson y Gerbing (1984), encontraron que, en modelos correctamente especificados, los valores impropios eran más posibles bajo algunas de las mismas condiciones que llevaban a la no convergencia en la estimación de los parámetros. Específicamente, muestras pequeñas y dos variables indicadoras por variable latente conducían a variancias negativas.

2.9.2.2 Prueba de hipótesis para los coeficientes

Los errores estándar de los estimadores presentados en (2.8), pueden ser empleados para probar si un parámetro es igual a cierta constante. Si $\hat{\theta}_i$ es un estimador de un parámetro, con θ_i como valor del parámetro según la hipótesis nula y $S_{\hat{\theta}}$ el error estándar asintótico del estimador, entonces se puede formar un estadístico de prueba como:

$$\frac{\hat{\theta}_i - \theta_i}{S_{\hat{\theta}}} \quad (2.60)$$

El cociente (2.60) se aproxima a un estadístico Z. Bajo la H_0 : $\theta_i = 0$, el estimador ML o GLS aproximadamente dos veces su error estándar asintótico se considera significativo al nivel de significación de 5%.

Una situación frecuente en la que los errores estándar no son fidedignos es cuando se analiza la matriz de correlación en lugar de la de covariancia (Boomsma, 1983).

2.9.2.3 Matriz de correlación asintótica de los estimadores de los parámetros

Una transformación de la matriz asintótica estimada de variancia-covariancia de los estimadores de los parámetros provee la matriz asintótica de correlación de los estimadores de los parámetros. La transformación se basa en la definición de la correlación como el cociente de la covariancia de dos variables y el producto de sus desviaciones estándar:

$$a\hat{\rho}_{\hat{\theta}_i, \hat{\theta}_j} = \frac{a \text{cov}(\hat{\theta}_i, \hat{\theta}_j)}{\sqrt{a \text{var}(\hat{\theta}_i) a \text{var}(\hat{\theta}_j)}} \quad (2.61)$$

Si (2.61) se aplica a todos los estimadores de los parámetros, resulta la matriz de correlación estimada. Correlaciones muy altas entre los estimadores de los parámetros son indicativas de severa multicolinealidad.

Del mismo modo que lo expresado en la sección anterior, la matriz asintótica de correlación es fidedigna en el análisis de matrices de covariancia pero no para las de correlación.

2.9.2.4 R^2 de las variables observadas

Otra medida del ajuste de los componentes es el coeficiente de correlación múltiple al cuadrado para cada variable x_i , que se calcula como:

$$R_{xi}^2 = 1 - \frac{\text{var}(\delta_i)}{\hat{\sigma}_{ii}} \quad (2.62)$$

donde $\hat{\sigma}_{ii}$ es la variancia de x_i predicha por el modelo. La expresión (2.62) se aplica, también a las variables observadas dependientes (y_i) y a las variables latentes endógenas (η). Una forma directa de obtener el R^2 es simplemente elevando al cuadrado la ponderación estandarizada (coeficiente estandarizado) de la variable.

La interpretación de esta medida es similar a la del análisis de regresión. El coeficiente R^2 es un indicador de la fiabilidad de una variable indicadora. Valores altos de R^2 indican que una alta proporción de la variancia de la variable es explicada por la variable latente.

2.9.2.5 Fiabilidad compuesta

El índice de fiabilidad compuesta refleja la consistencia interna de las variables indicadoras que miden una variable latente dada y está expresado como:

$$Fiabilidad\ compuesta = \frac{(\sum \hat{\lambda}_i)^2}{(\sum \hat{\lambda}_i)^2 + \sum (1 - \hat{\lambda}_i^2)} \quad (2.63)$$

donde $\hat{\lambda}_i$ son las ponderaciones estandarizadas para la variable latente; las variancias de error asociadas con las variables indicadoras se calculan como $1 - (\hat{\lambda}_i)^2$.

Se considera que valores entre 0.6 y 0.7 son valores mínimos aceptables (Hatcher, 1994).

2.9.2.6 Variancia extraída

Fornell y Larcker (1981) presentan otro índice de fiabilidad llamado estimador de la variancia extraída. Esta medida refleja la cantidad total de la variancia de los indicadores tenida en cuenta por la variable latente en relación a la cantidad de variancia debida al error de medición. Los mayores valores de la variancia extraída se producen cuando los indicadores son verdaderamente representativos de la variable latente. Se calcula como:

$$Variancia\ extraída = \frac{\sum \hat{\lambda}_i^2}{\sum \hat{\lambda}_i^2 + \sum (1 - \hat{\lambda}_i^2)} \quad (2.64)$$

Este índice es muy parecido a la medida de la fiabilidad compuesta pero difiere en que las ponderaciones estandarizadas se elevan al cuadrado antes de sumarlas. En general se sugiere que el valor de la variancia extraída debería exceder 0.50 para una variable latente (Fornell y Larcker, 1981).

2.9.3 Comparación de modelos

Frecuentemente, se ajusta más de un modelo a los mismos datos, ya sea modelos con muy diferente estructura o con diferencias menores entre uno o dos parámetros. Una de las estrategias para realizar esto es comparar los valores de los índices de ajuste global, siendo preferibles, en este caso, las medidas que tienen en cuenta los grados de libertad. Sin embargo, esta comparación no deja de ser descriptiva, estas medidas no permiten probar hipótesis estadísticas.

Si se opta por probar hipótesis para la comparación de modelos, existen tres pruebas que son aplicables a modelos anidados:

- (1) Prueba del cociente de verosimilitud (LR) o diferencia de χ^2 ,
- (2) Prueba del multiplicador de Lagrange (LM), y,
- (3) Prueba de Wald (W).

2.9.3.1 Prueba del cociente de verosimilitud (LR)

Sean $\hat{\theta}_r$ el estimador ML para el modelo anidado restringido, y $\hat{\theta}_u$ el correspondiente al modelo sin la restricción. El estadístico LR es:

$$LR = -2[\ln L(\hat{\theta}_r) - \ln L(\hat{\theta}_u)] \quad (2.65)$$

que tiene una distribución χ^2 cuando el modelo restringido es válido. En la práctica, se calcula el estadístico LR como la diferencia entre los estimadores χ^2 usuales para los modelos con restricción y sin restricción, con el número de grados de libertad igual a la diferencia de los grados de libertad de los dos modelos. Por esta razón, LR es comúnmente conocido como la prueba de diferencia de χ^2 . La expresión (2.65) es una generalización de (2.46).

Una desventaja de esta prueba a nivel computacional es que dos modelos deben estimarse para cada par de comparaciones. Esta limitante no existe en las pruebas del multiplicador de Lagrange y Wald.

2.9.3.2 Prueba del Multiplicador de Lagrange (LM)

La prueba LM compara el ajuste de modelos restrictivos a menos restrictivos y sólo requiere la estimación del modelo restrictivo. El estadístico LM está basado en $\partial \ln L(\theta) / \partial \theta$, donde $\ln L(\theta)$ es el logaritmo de la función de verosimilitud del modelo no restrictivo. Esta derivada parcial denominada *puntaje* $s(\theta)$, es el cambio en el logaritmo de verosimilitud por cambios en θ . Los elementos no restrictivos de $\hat{\theta}_r$ deberían tener una derivada parcial nula, ya que el estimador ML resuelve para $\hat{\theta}_r$ llevando las derivadas parciales a 0. Sin embargo, los elementos de $s(\theta)$ correspondientes a los parámetros restringidos solo son ceros cuando las restricciones se cumplen exactamente. Sustituyendo $\hat{\theta}_r$ por θ en $s(\theta)$, el examen de las magnitudes de las desviaciones del 0 para los parámetros restrictivos indica la validez de las restricciones. Debido a que las fluctuaciones en el muestreo llevan a valores apartados de cero, aún cuando la restricción es válida, se requiere tener en cuenta la variabilidad. En este sentido, el estadístico de prueba empleado es:

$$LM = [s(\hat{\theta}_r)]' \mathbf{I}^{-1}(\hat{\theta}_r) s(\hat{\theta}_r) \quad (2.66)$$

donde $\mathbf{I}^{-1}(\hat{\theta}_r) = (-E[\partial^2 \ln L(\theta) / \partial \theta \partial \theta'])^{-1}$ evaluada en $\hat{\theta}_r$ y $s(\hat{\theta}_r)$ es $s(\theta)$ evaluada en $\hat{\theta}_r$. La distribución de LM es una χ^2 con grados de libertad igual a la diferencia de

grados de libertad de los dos modelos, el restrictivo y el no restrictivo. Dada la relación entre F_{ML} y el $\ln L(\boldsymbol{\theta})$, el estadístico de prueba puede ser escrito como:

$$LM = \frac{N-1}{2} \left(\frac{\partial F_{ML}}{\partial \boldsymbol{\theta}} \right)' \left[E \left(\frac{\partial^2 F_{ML}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right) \right]^{-1} \left(\frac{\partial F_{ML}}{\partial \boldsymbol{\theta}} \right) \quad (2.67)$$

donde $\boldsymbol{\theta}$ es evaluado en $\hat{\boldsymbol{\theta}}_r$.

2.9.3.3 Prueba de Wald (W)

La prueba de Wald (W) determina el grado en que $\hat{\boldsymbol{\theta}}_u$ se aparta de las restricciones impuestas en el modelo anidado. Sea $r(\boldsymbol{\theta})$ un vector ($r \times 1$), siendo r menor a la dimensión fila de $\boldsymbol{\theta}$, que contiene las restricciones en un modelo no restringido. El estadístico de prueba es

$$W = \left[r(\hat{\boldsymbol{\theta}}_u) \right]' \left\{ \left[\frac{\partial r(\hat{\boldsymbol{\theta}}_u)}{\partial \hat{\boldsymbol{\theta}}_u} \right] \left[a \text{cov}(\hat{\boldsymbol{\theta}}_u) \right] \left[\frac{\partial r(\hat{\boldsymbol{\theta}}_u)}{\partial \hat{\boldsymbol{\theta}}_u} \right]' \right\}^{-1} \left[r(\hat{\boldsymbol{\theta}}_u) \right] \quad (2.68)$$

donde $r(\hat{\boldsymbol{\theta}}_u)$ es $r(\boldsymbol{\theta}_u)$ evaluado en $\hat{\boldsymbol{\theta}}_u$ y $[a \text{cov}(\hat{\boldsymbol{\theta}}_u)]$ es un estimador de la matriz de covariancia asintótica de $\hat{\boldsymbol{\theta}}_u$. El término entre llaves $\{ \cdot \}$ es un estimador de la matriz de covariancia asintótica de $r(\hat{\boldsymbol{\theta}}_u)$. De este modo, W es la inversa de la matriz de covariancia asintótica de las restricciones impuestas por el modelo restrictivo, pre y postmultiplicada por $r(\boldsymbol{\theta}_u)$ evaluada en $\hat{\boldsymbol{\theta}}_u$. Asintóticamente W sigue una distribución χ^2 con grados de libertad iguales al número de restricciones en $r(\boldsymbol{\theta}_u)$ cuando el modelo restrictivo es válido.

Cuando $r(\boldsymbol{\theta}_u)$ consiste en una sola restricción de $\theta_1 = 0$, W es

$$W = \frac{\hat{\theta}_1^2}{a \text{var}(\hat{\theta}_1)} \quad (2.69)$$

donde $a \text{var}(\hat{\theta})$ es la variancia asintótica estimada de $\hat{\theta}_1$ y W tiene 1 grado de libertad. La ecuación (2.69) es el cuadrado de la variable Normal estándar, Z , para probar la hipótesis de un parámetro. Es decir que se puede ver a la prueba de Wald como una generalización de la prueba de Z usual.

2.9.3.4 Medida de ajuste de la componente estructural del modelo

Como se ha mencionado, un modelo teórico consiste de dos componentes, un modelo de medida que describe la relación entre las variables latentes y sus variables indicadoras, y un modelo estructural que describe las relaciones causales de las variables latentes entre sí. Los índices de bondad de ajuste que se describieron hasta aquí reflejan el ajuste global del modelo de medida y el estructural combinados. Sin embargo, la mayor parte de los modelos teóricos contienen un pequeño número de variables latentes y un gran número de variables indicadoras. Esto significa que la porción correspondiente al modelo de medida, generalmente, contiene muchos más parámetros a ser estimados que la porción estructural. En consecuencia, los índices de ajuste global, tales como NNFI y CFI, están mucho más influenciados por el ajuste de la porción del modelo de medida que del ajuste de la porción estructural. De este modo, podría ocurrir que la porción correspondiente al modelo de medida presentara un muy buen ajuste por lo que las medidas de ajuste global fueran muy buenas, aún cuando exista una mala especificación en la porción estructural del modelo. Como una solución a este problema, Mulaik *et al.* (1989), proponen un índice de ajuste relativo normado (RNFI), que refleja el ajuste sólo de la porción estructural del modelo, y no es influenciado por el ajuste del modelo de medida. Su interpretación es similar a la del NFI, en que valores altos, cercanos a 1, indican que las relaciones causales entre las variables latentes establecidas en la hipótesis proveen un buen ajuste a los datos. Se puede calcular el RNFI como sigue:

$$RNFI = \frac{\chi_u^2 - \chi_j^2}{\chi_u^2 - \chi_m^2 - (df_j - df_m)} \quad (2.70)$$

donde

$\chi_u^2 = \chi^2$ del modelo con variables latentes no correlacionadas

$\chi_j^2 = \chi^2$ del modelo de interés

$\chi_m^2 = \chi^2$ del modelo de medida

df_j = grados de libertad para el modelo de interés

df_m = grados de libertad para el modelo de medida

El modelo con variables latentes no correlacionadas predice que las variables latentes no se correlacionan entre sí.

Capítulo 3: ESPECIFICACION DEL MODELO ESTRUCTURAL PARA LA VALIDACIÓN DE LA EVALUACIÓN DE VACUNAS CONTRA BoHV-1

3.1. Variables del modelo

En el desarrollo de un modelo para la evaluación de la potencia de una vacuna para BoVH1 se busca verificar la concordancia de la respuesta de anticuerpos en cobayos con la potencia de la vacuna en bovinos (respuesta de anticuerpos en bovinos) y relacionar ambos datos serológicos con la protección a la enfermedad y a la infección en bovinos (eficacia de la vacuna). Este sistema involucra variables latentes ya que las variables protección a la enfermedad, protección a la infección y la calidad de vacuna no son directamente observables.

La situación ideal es estimar el grado de **protección a la infección y a la enfermedad** evaluando **la calidad de la vacuna** a partir del nivel de anticuerpos alcanzado por los bovinos y, mejor aún, por los cobayos.

Un animal está protegido contra la infección si resulta refractario contra la infección. En estos casos se dice que la inmunidad conferida por la vacuna es esterilizante ya que no permite el ingreso del virus al organismo. No existen, actualmente, vacunas contra BoVH1 que protejan contra la infección en sentido estricto, sino que se debe considerar a este tipo de protección desde la perspectiva de distintos niveles de protección.

Un animal protegido contra la enfermedad es aquel animal vacunado que al estar expuesto al virus se infecta, el agente se multiplica en su organismo y se establece la infección latente y puede ser transmisor del virus. Sin embargo, no presenta síntomas de la enfermedad.

El concepto fundamental en este trabajo es la calidad de la vacuna contra el virus BoHV-1, causante de la Rinotraqueítis infecciosa bovina. Este concepto presenta tres dimensiones que originan tres variables latentes. La protección a la infección y la calidad asignada por el laboratorio (ξ_1 y ξ_2 , respectivamente) son dos variables latentes exógenas y la protección a la enfermedad (η) es una variable latente endógena. La razón para clasificar estas variables en exógenas y endógena se basa en que no existe en el sistema una variable que pueda ser causa de la protección a la infección. Por otra parte, si un animal está protegido contra la infección estará protegido contra la enfermedad. La enfermedad es causada por la infección, pero no es posible el caso contrario. La calidad asignada por el laboratorio estará relacionada con la protección a la infección, pero ninguna de las dos es causa de la otra. Para comprender este punto de vista se debe tener presente que la unidad de análisis es el lote de vacuna y no el animal, y que, como se desea analizar a las evaluaciones serológicas como evaluaciones alternativas a la protección, se parte de que los animales que participan en el desafío no son los mismos que son evaluados en el laboratorio. Estos últimos animales no han sido desafiados con el virus.

Se consideran 4 variables indicadoras de la protección a la infección:

X₁: cantidad máxima de virus excretado en un día o pico de excreción en DICT50/ml (dosis infecciosas cultivo de tejido 50% / ml)

X₂: tiempo en días al inicio de excreción de virus a partir de la descarga

X₃: duración en días de la excreción de virus.

X₄: área bajo la curva obtenida por la cantidad de virus excretado durante 14 días a partir de la descarga.

Estas variables son consideradas indicadoras de la protección a la infección porque evalúan el nivel de virus que tiene el animal. Asimismo, puede observarse que la protección de un animal implica una menor probabilidad de contagio del virus. Si el animal excreta virus durante menos tiempo o si la cantidad de virus excretado es menor, el animal tendrá un mayor nivel de protección a la infección y la probabilidad de que los otros animales del rodeo estén en contacto con el virus será menor. En este sentido, el área bajo la curva de la cantidad de virus excretado en el tiempo, que tiene como antecedente el trabajo de Parreño *et al.* (2010a), se presenta como una variable resumen.

Como indicadoras de la protección a la enfermedad se consideran 5 variables observadas o manifiestas:

Y₁: presencia y severidad de la rinitis (descarga nasal con tipo de secreción, serosa, seromucosa, mucopurulenta) medida como puntaje.

Y₂: el tiempo desde la descarga hasta el pico de severidad en días

Y₃: duración de síntomas en días

Y₄: temperatura del animal (total de días con hipertermia).

Y₅: área bajo la curva obtenida por el puntaje de severidad durante 14 días a partir de la descarga.

La protección a la enfermedad interesa, básicamente, desde el punto de vista económico de la producción animal. Si el animal está protegido contra la enfermedad no desarrollará (o lo hará en menor medida) los signos y síntomas de la enfermedad, aún siendo portador sano del virus. La ausencia de los síntomas de la Rinotraqueítis infecciosa bovina permite que el animal no se encuentre afectado en su producción. Del mismo modo que con la variable protección a la infección, en este caso, también, se pueden considerar niveles de protección a la enfermedad.

Para la representación del constructo calidad de vacuna asignada por el laboratorio se emplean como indicadoras 4 variables: los títulos de anticuerpos medidos por ELISA y seroneutralización en bovinos y cobayos:

X₅: Título de Ac. por ELISA en bovinos

X₆: Título de Ac. por seroneutralización en bovinos

X₇: Título de Ac. por ELISA en cobayos

X₈: Título de Ac. por seroneutralización en cobayos

La técnica ELISA (*Enzyme-Linked ImmunoSorbent Assay*) se basa en la detección de un antígeno inmovilizado sobre una fase sólida mediante anticuerpos que directa o indirectamente producen una reacción cuyo producto, por ejemplo un colorante, puede ser medido por espectrofotometría. La neutralización viral (SN) se define como la pérdida de la infectividad de un virus, por la reacción del mismo con un anticuerpo específico. Si bien varias pruebas ELISA fueron desarrolladas para

determinar los anticuerpos para BoHV-1 y demostraron ser más sensibles y específicas que la prueba de seroneutralización (OIE, 2008b; Kramps *et al.*, 2004), esta última sigue considerándose como la técnica “*gold standard*” para su empleo en pruebas de potencia de vacunas (OIE, 2008; CFR, 1985).

La protección a la infección no sólo es de interés por su relación causal con la protección a la enfermedad sino, también, porque, como se mencionó antes, un animal adecuadamente protegido contra la infección tiene menor probabilidad de diseminar el virus al rodeo. La calidad de la vacuna asignada por el laboratorio es, en realidad, una forma alternativa de evaluar la vacuna sin tener que desafiar a los animales destino con el virus. Por esta razón, de comprobarse una relación fuerte con la protección, se constituye en un método económico para calificar a una vacuna. Por otra parte, si bien los cobayos no desarrollan la enfermedad, producen anticuerpos para el virus BoHV-1 al ser vacunados.

3.2. El modelo de ecuaciones estructurales para la validación de un método de evaluación de vacunas

Para la modelización de las relaciones de las variables involucradas en la evaluación de la vacuna bovina se consideran los dos componentes de las ecuaciones estructurales. El primer componente es el modelo de variable latente:

$$\boldsymbol{\eta} = \boldsymbol{\gamma}_1 \boldsymbol{\xi}_1 + \boldsymbol{\zeta} \quad (3.1)$$

donde $\boldsymbol{\eta}$ es la variable latente aleatoria endógena, que en este caso es un escalar (η = protección a la enfermedad); $\boldsymbol{\xi}_1$, es la variable latente aleatoria exógena protección a la infección; $\boldsymbol{\gamma}_1$ es el coeficiente para los efectos de $\boldsymbol{\xi}_1$ en $\boldsymbol{\eta}$. El disturbio $\boldsymbol{\zeta}$ se asume que tiene esperanza igual a cero [$E(\boldsymbol{\zeta}) = \mathbf{0}$] y que no está correlacionado con $\boldsymbol{\xi}_1$.

El segundo componente del sistema general es el modelo de medida:

$$\mathbf{y} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon} \quad (3.2)$$

$$\mathbf{x} = \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta} \quad (3.3)$$

donde

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} \quad \boldsymbol{\Lambda}_y = \begin{bmatrix} \lambda_{10} \\ \lambda_{11} \\ \lambda_{12} \\ \lambda_{13} \\ \lambda_{14} \end{bmatrix} \quad \eta = [\eta_1] \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{bmatrix} \quad (3.4)$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{bmatrix} \quad \Lambda_x = \begin{bmatrix} \lambda_1 & 0 \\ \lambda_2 & 0 \\ \lambda_3 & 0 \\ \lambda_4 & 0 \\ 0 & \lambda_5 \\ 0 & \lambda_6 \\ 0 & \lambda_7 \\ 0 & \lambda_8 \end{bmatrix} \quad \xi = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \quad \delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ \delta_7 \\ \delta_8 \end{bmatrix} \quad (3.5)$$

Los \mathbf{y} (5x1) y \mathbf{x} (8x1) son vectores de variables observadas, Λ_y (5x1) y Λ_x (8x2) son las matrices de coeficientes que muestran la relación de \mathbf{y} con $\boldsymbol{\eta}$ y de \mathbf{x} con $\boldsymbol{\xi}$, respectivamente, y $\boldsymbol{\varepsilon}$ (5x1) y $\boldsymbol{\delta}$ (8x1) son los errores de medición para \mathbf{y} y \mathbf{x} , respectivamente. Los errores de medición se suponen no correlacionados con $\boldsymbol{\xi}$ y $\boldsymbol{\zeta}$ y entre sí. Las esperanzas de $\boldsymbol{\varepsilon}$ y $\boldsymbol{\delta}$ son iguales a cero. $\boldsymbol{\eta}$, $\boldsymbol{\xi}$, \mathbf{y} y \mathbf{x} se escriben como desviaciones de sus medias. Es necesario considerar un modelo alternativo que sea menos restrictivo en el establecimiento de la no correlación entre errores para las variables indicadoras de la calidad de la vacuna por laboratorio. Se debe tener en cuenta que las determinaciones de anticuerpos por ELISA y seroneutralización se realizan sobre el mismo animal bovino o cobayo.

Para emplear el modelo general se debe especificar el patrón de elementos en cada una de las matrices Λ_y , Λ_x , Φ , Ψ , Θ_ε y Θ_δ --

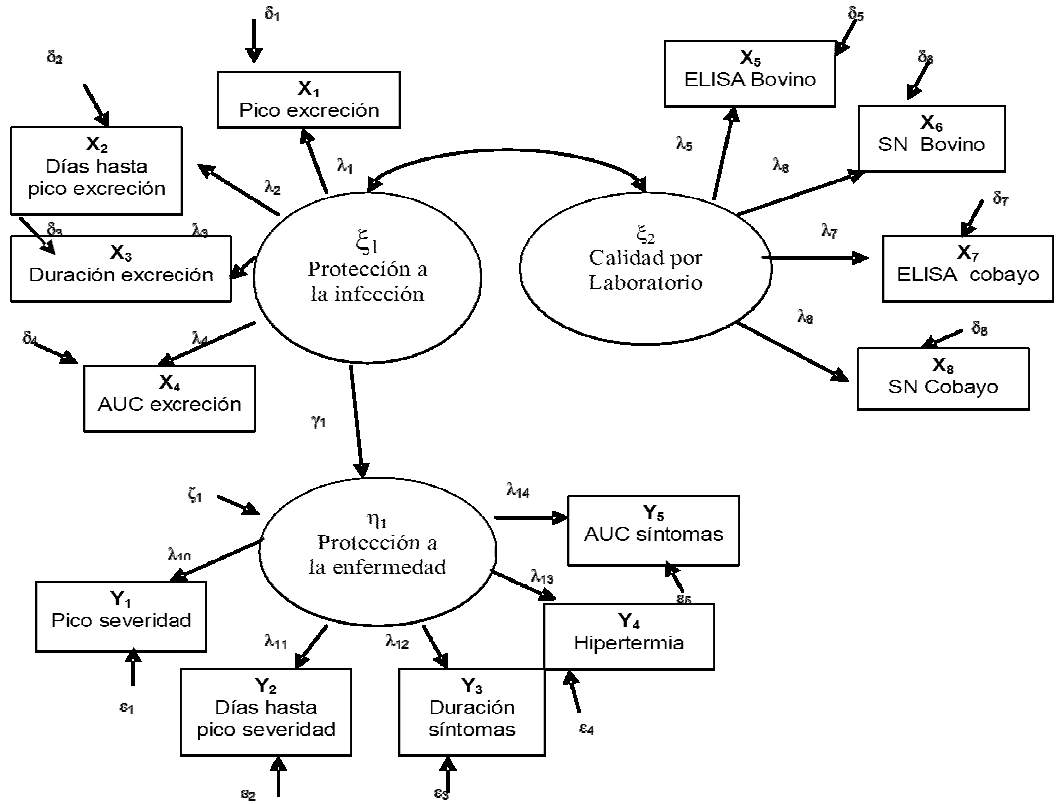
La $E(\boldsymbol{\xi}\boldsymbol{\xi}')$ constituye la matriz de covariancia de las variables exógenas $\boldsymbol{\xi}$, y se simboliza con la letra griega Φ . La $E(\boldsymbol{\zeta}\boldsymbol{\zeta}')$, Ψ , en este caso es un escalar ya que hay una sola variable latente endógena siendo ψ_{11} la variancia de la variable η_1 que no está explicada por las variables explicativas incluidas en la ecuación. Θ_ε es la $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')$ (matriz de covarianza de $\boldsymbol{\varepsilon}$) y Θ_δ es la $E(\boldsymbol{\delta}\boldsymbol{\delta}')$ (matriz de covarianza de $\boldsymbol{\delta}$). Θ_ε y Θ_δ son matrices diagonales, con las variancias de los errores correspondientes en la diagonal principal. En el caso del modelo alternativo las covarianzas entre los errores ε_6 , y ε_8 no son restringidas a 0.

La matriz de covariancia de las variables latentes exógenas, Φ , puede escribirse como:

$$\Phi = \begin{bmatrix} \phi_{11} & \\ \phi_{21} & \phi_{22} \end{bmatrix} \quad (3.6)$$

Las ecuaciones (3.1), (3.2) y (3.3) pueden representarse gráficamente mediante el diagrama de senderos (Figura 3.1).

Figura 3.1: Diagrama de senderos de las relaciones entre variables en la evaluación de vacunas contra el virus BoHV-1



La matriz de covariancia de las variables observadas, Σ está formada por 4 submatrices según (3.7)

$$\Sigma = \begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix} \quad (3.7)$$

donde

$$\Sigma_{yy} = \begin{bmatrix} \sigma_{y_1}^2 & \sigma_{y_2 y_1} & \sigma_{y_3 y_1} & \sigma_{y_4 y_1} & \sigma_{y_5 y_1} \\ \sigma_{y_2 y_1} & \sigma_{y_2}^2 & \sigma_{y_3 y_2} & \sigma_{y_4 y_2} & \sigma_{y_5 y_2} \\ \sigma_{y_3 y_1} & \sigma_{y_3 y_2} & \sigma_{y_3}^2 & \sigma_{y_4 y_3} & \sigma_{y_5 y_3} \\ \sigma_{y_4 y_1} & \sigma_{y_4 y_2} & \sigma_{y_4 y_3} & \sigma_{y_4}^2 & \sigma_{y_5 y_4} \\ \sigma_{y_5 y_1} & \sigma_{y_5 y_2} & \sigma_{y_5 y_3} & \sigma_{y_5 y_4} & \sigma_{y_5}^2 \end{bmatrix} \quad (3.8)$$

$$\Sigma_{xx} = \begin{bmatrix} \sigma_{x_1}^2 & & & & & & & & \\ \sigma_{x_2 x_1} & \sigma_{x_2}^2 & & & & & & & \\ \sigma_{x_3 x_1} & \sigma_{x_3 x_2} & \sigma_{x_3}^2 & & & & & & \\ \sigma_{x_4 x_1} & \sigma_{x_4 x_2} & \sigma_{x_4 x_3} & \sigma_{x_4}^2 & & & & & \\ \sigma_{x_5 x_1} & \sigma_{x_5 x_2} & \sigma_{x_5 x_3} & \sigma_{x_5 x_4} & \sigma_{x_5}^2 & & & & \\ \sigma_{x_6 x_1} & \sigma_{x_6 x_2} & \sigma_{x_6 x_3} & \sigma_{x_6 x_4} & \sigma_{x_6 x_5} & \sigma_{x_6}^2 & & & \\ \sigma_{x_7 x_1} & \sigma_{x_7 x_2} & \sigma_{x_7 x_3} & \sigma_{x_7 x_4} & \sigma_{x_7 x_5} & \sigma_{x_7 x_6} & \sigma_{x_7}^2 & & \\ \sigma_{x_8 x_1} & \sigma_{x_8 x_2} & \sigma_{x_8 x_3} & \sigma_{x_8 x_4} & \sigma_{x_8 x_5} & \sigma_{x_8 x_6} & \sigma_{x_8 x_7} & \sigma_{x_8}^2 \end{bmatrix} \quad (3.9)$$

$$\Sigma_{xy} = \begin{bmatrix} \sigma_{x_1 y_1} & \sigma_{x_1 y_2} & \sigma_{x_1 y_3} & \sigma_{x_1 y_4} & \sigma_{x_1 y_5} \\ \sigma_{x_2 y_1} & \sigma_{x_2 y_2} & \sigma_{x_2 y_3} & \sigma_{x_2 y_4} & \sigma_{x_2 y_5} \\ \sigma_{x_3 y_1} & \sigma_{x_3 y_2} & \sigma_{x_3 y_3} & \sigma_{x_3 y_4} & \sigma_{x_3 y_5} \\ \sigma_{x_4 y_1} & \sigma_{x_4 y_2} & \sigma_{x_4 y_3} & \sigma_{x_4 y_4} & \sigma_{x_4 y_5} \\ \sigma_{x_5 y_1} & \sigma_{x_5 y_2} & \sigma_{x_5 y_3} & \sigma_{x_5 y_4} & \sigma_{x_5 y_5} \\ \sigma_{x_6 y_1} & \sigma_{x_6 y_2} & \sigma_{x_6 y_3} & \sigma_{x_6 y_4} & \sigma_{x_6 y_5} \\ \sigma_{x_7 y_1} & \sigma_{x_7 y_2} & \sigma_{x_7 y_3} & \sigma_{x_7 y_4} & \sigma_{x_7 y_5} \\ \sigma_{x_8 y_1} & \sigma_{x_8 y_2} & \sigma_{x_8 y_3} & \sigma_{x_8 y_4} & \sigma_{x_8 y_5} \end{bmatrix} \quad (3.10)$$

$$y \Sigma_{yx} = \Sigma_{xy}'$$

Como se demostró en la sección (2.4) del capítulo anterior, se puede escribir la matriz de covariancia como una función de los parámetros del modelo. De este modo se tiene que:

$$\begin{aligned} \Sigma_{yy}(\theta) &= E(\mathbf{y}\mathbf{y}') \\ &= E[(\Lambda_y \eta_1 + \epsilon)(\eta_1' \Lambda_y' + \epsilon')] \\ &= \Lambda_y (\gamma_1^2 \phi_1 + \psi_1) \Lambda_y' + \Theta_\epsilon \end{aligned} \quad (3.11)$$

$$\begin{aligned} \Sigma_{yx}(\theta) &= E(\mathbf{y}\mathbf{x}') \\ &= E[(\Lambda_y \eta_1 + \epsilon)(\Lambda_x \xi + \delta)'] \\ &= \Lambda_y E(\eta_1 \xi') \Lambda_x' \\ &= \Lambda_y E[(\gamma_{11} \zeta_1 + \zeta_1) \xi'] \Lambda_x' \\ &= \Lambda_y \gamma_{11} (\Phi_{11} \Phi_{12}) \Lambda_x' \end{aligned} \quad (3.12)$$

$$\begin{aligned} \Sigma_{xx}(\theta) &= E(\mathbf{x}\mathbf{x}') \\ &= E[(\Lambda_x \xi + \delta)(\xi' \Lambda_x' + \delta')] \\ &= \Lambda_x E(\xi \xi') \Lambda_x' + \Theta_\delta \\ &= \Lambda_x \Phi \Lambda_x' + \Theta_\delta \end{aligned} \quad (3.13)$$

Escribiendo (3.11), (3.12) y (3.13) con el detalle de sus elementos se llega a (3.14), (3.15) y (3.16), respectivamente.

3.3. Experimentos y medición de variables

La aplicación del modelo presentado en la sección 3.2 se basó en una serie de experimentos que llevó a cabo un grupo de investigación del Instituto de Virología del INTA de Castelar (Parreño *et al.*, 2010a). Este estudio incluyó un total de 553 novillos de ambos géneros de entre 6 y 12 meses de edad de las razas Aberdeen Angus, Hereford y sus cruza. Los ensayos de vacunación fueron conducidos en 12 establecimientos de la provincia de Buenos Aires, Argentina, seleccionándose rodeos sin historia previa de vacunación contra el virus. Como la infección por BoHV-1 es endémica en Argentina, las vacunas fueron evaluadas en animales seronegativos de rodeos libres del virus, y también en bovinos seronegativos y seropositivos de rodeos endémicos para considerar la variabilidad de la población real. En los ensayos conducidos en establecimientos endémicos, los animales positivos y negativos fueron distribuidos aleatoriamente de modo de obtener valores medios y variancias similares de títulos de anticuerpos (Ab) al inicio del estudio. En cada ensayo, los bovinos fueron vacunados con dos dosis de vacuna distanciadas 30 días entre sí según recomendación del marbete de cada una de las 22 vacunas comerciales probadas, por vía subcutánea con dosis de 3 a 5 ml. Las vacunas formuladas especialmente para la validación del modelo, llamadas “vacunas de referencia”, fueron aplicadas siguiendo las mismas dosis e intervalos de tiempo que las vacunas comerciales. Las muestras de sangre para la extracción de suero fueron obtenidas por punción de la vena yugular o coccígea, a los 60 días post-vacunación (dpv). Se incluyeron placebos y animales no vacunados como controles que fueron expuestos a condiciones ambientales similares a las de los bovinos vacunados. Sólo se consideró como válido para el análisis el ensayo en el que no existió seroconversión en los bovinos controles. La mayoría de los experimentos fueron llevados a cabo como ciegos (veterinario, laboratorio).

Para el desarrollo del modelo de animal de laboratorio se emplearon 497 cobayos (*Cavia porcellus*), SiS Al, de alrededor de 350-400 g de peso. Los animales fueron vacunados con dosis distanciadas 21 días, por la vía intramuscular o pata trasera. La dosis administrada a los cobayos fue 1/5 del volumen de la dosis dada a los bovinos. La extracción de sangre se realizó a los 30 días posvacunación vía punción cardíaca bajo anestesia, siguiendo las recomendaciones de ECVAM para bienestar animal (Leenaars *et al.*, 1999).

Las evaluaciones del título de anticuerpos se realizaron empleando las técnicas de seroneutralización (SN) y un ELISA indirecto tal como se describe en Del Medico *et al.* (2006) y Romera *et al.* (2000). El ensayo de neutralización de virus usado en este estudio es el considerado como la prueba tradicional o “Gold Standard”. Las muestras seronegativas recibieron el valor arbitrario de 0.30. El título de anticuerpos de cada muestra fue en \log_{10} del recíproco de la dilución máxima de suero con una densidad óptica corregida OD_{405c}. Los resultados de ELISA fueron normalizados expresando los valores corregidos de OD como porcentajes de un único control de suero altamente positivo (PP%) incluido en cada plato y en cada corrida de ELISA. La validación de los ensayos cumplió las normas ISO/IEC 17025.

Para estudiar la relación entre la dosis y la respuesta serológica, se formularon tres conjuntos de vacunas oleosas, del tipo agua/aceite, que contenían concentraciones crecientes de BoHV-1, cubriendo el rango desde 10^5 a 10^7 TCID₅₀/dosis, y dos vacunas

con 10^8 TCID₅₀/dosis, junto con concentraciones fijas del virus de parainfluenza bovina tipo 3 y el virus de diarrea viral del bovino. Los antígenos virales fueron inactivados mediante etilenamina binaria y las vacunas fueron formuladas en un equipo piloto a escala. El adyuvante empleado en la formulación consistió en una mezcla de 0.67% de Polisorbato 80, 2.1% de Sorbitan 80 Monooleato y 57.9% de aceite mineral en una proporción aceite:agua de 60:40. Cada conjunto de vacunas fue formulado con una diferente dilución del mismo lote de antígeno que fue previamente diluido o concentrado de modo de obtener la concentración de antígeno (Ag) deseada. Estas vacunas de concentración conocida de antígeno se mencionan como “vacunas de referencia” y fueron empleadas para estimar la curva de calibración dosis-respuesta en bovinos y cobayos. Cada conjunto de vacunas de referencia fue probado en dos experimentos independientes de bovinos y cobayos.

Los ensayos de desafío, en los que los bovinos fueron inoculados con el virus, se llevaron a cabo en el CICV como se describió en Romera *et al.* (2000). Bovinos seronegativos fueron vacunados con vacunas que habían sido clasificadas como “muy satisfactorias” o “satisfactorias” de acuerdo a la prueba serológica de bovinos y cobayos (López *et al.*, 2007; Parreño *et al.*, 2010a). Los animales que recibieron el placebo fueron considerados como inoculados con una vacuna “no satisfactoria”. Grupos de 6 bovinos fueron vacunados con dos dosis de cada vacuna, distanciadas 30 días. A los 90 dpv los animales fueron desafiados por vía intranasal con el virus BoHV-1 a una concentración de $10^{7.5}$ TCID₅₀/ml. Los animales fueron monitoreados por infección y desarrollo de la enfermedad. La infección del virus se midió por la duración y el pico excretorio del virus y por el área bajo la curva de la excreción durante 14 días luego del desafío (AUCi). La enfermedad fue evaluada por la duración y severidad de los signos clínicos de IBR y por el área bajo la curva de la presencia de los signos clínicos durante los 14 días posteriores al desafío (AUCs). Ambas áreas bajo las curvas fueron calculadas empleando el software MedCalc® versión 11.1.1.0. En el Cuadro 3.1 se presentan las medidas descriptivas de las variables medidas.

Todos los ensayos descriptos hasta aquí fueron llevados a cabo de acuerdo a un diseño que no contemplaba un análisis multivariado, y menos aún, el análisis de un modelo de ecuaciones estructurales. Por esta razón, se presentaron algunas dificultades para el empleo de la base de datos original en este trabajo. Por un lado, fue bajo el número de lotes de vacunas, que, además de ser empleadas en los ensayos de dosis-respuesta (evaluación serológica), también fueran inoculados en los ensayos de desafío. Por otro lado, un grupo de bovinos proveniente de uno de los campos que participó del ensayo de desafío no respondió a vacunas con alta carga de antígenos, comportándose de modo distinto al resto de los bovinos vacunados en los ensayos de dosis-respuesta, y también a los cobayos que sí siguieron el patrón de respuesta de esos ensayos. Esta situación resultó especialmente seria debido al bajo número de animales que participó del desafío. Un tercer inconveniente fue que la variable *hipertermia* o *días con hipertermia*, contemplada en el modelo teórico como variable indicadora de la protección a la enfermedad, no fue medida en los ensayos de desafío.

La base de datos original se dividió en dos partes de modo de analizar la relación de las variables serológicas en cobayos y los bovinos del campo 3 (animales no respondedores) y el resto. En el Cuadro 3.2 se presentan los resultados del análisis de regresión para SN y ELISA. En este cuadro se agrega el análisis de regresión de las respuestas serológicas de bovinos y cobayos que no corresponden a los ensayos con desafío. Los resultados del coeficiente de determinación fueron bajos en el análisis de los animales del Campo 3, contrastando con los altos valores obtenidos, tanto en el resto de los animales ensayados junto con el desafío como en los ensayos para dosis-

respuesta. Asimismo, se observa que las pendientes de estos dos últimos grupos son similares.

Cuadro 3.1.: Medidas descriptivas de las variables manifiestas en la base de datos original

Variable	Número de animales	Media (Desvío estándar)	Mediana (Mínimo-Máximo)
X₁ Pico excretorio (DICT50/ml)	46	5.72 (1.77)	6.40 (0.00-7.50)
X₂ Días hasta el pico excretorio	46	5.30 (1.40)	5.00 (0.00-8.00)
X₃ Duración de excreción del virus (días)	46	5.28 (2.17)	5.00 (0.00-10.00)
X₄ Área bajo la curva de excreción de virus	46	27.97 (13.36)	25.78 (0.00-56.40)
Y₁ Pico de severidad	46	2.67 (0.82)	3.00 (1.00-4.00)
Y₂ Tiempo hasta pico de severidad (días)	46	6.24 (3.37)	6.00 (2.00-16.00)
Y₃ Duración de síntomas (días)	46	9.61 (3.83)	8.50 (4.00-17.00)
Y₄ Puntaje a los 14 días	46	10.74 (4.13)	10.00 (2.00-21.00)
Y₅ Área bajo la curva de severidad	46	19.85 (9.72)	19.00 (4.50-42.00)
X₅ Título de Ac por ELISA en bovino	51	2.20 (1.14)	2.20 (0.60-4.01)
X₆ Título de Ac por SN en bovino	43	1.55 (0.80)	1.69 (0.44-2.86)
X₇ Título de Ac por ELISA en cobayo	43	1.87 (0.88)	2.10 (0.60-3.00)
X₈ Título de Ac por SN en cobayo	45	2.68 (1.24)	3.01 (0.60-4.60)

Cuadro 3.2: Resultados del análisis de regresión de las variables serológicas de los cobayos en función de bovinos del Campo 3 (animales no respondedores) y del resto

Animales ensayados	Campo	ELISA				SN			
		b₀⁽¹⁾	b₁⁽²⁾	CME⁽³⁾	R²	b₀	b₁	CME⁽³⁾	R²
Ensayos con desafío	Campo 3 (n = 20)	2.06	0.62	0.21	0.40	1.47	0.48	0.15	0.35
	Resto (n = 25)	0.14	0.82	0.20	0.87	-0.22	1.11	0.03	0.98
Ensayos sin desafío (n=58)		-0.43	1.15	0.46	0.77	-0.04	1.10	0.08	0.90

(1) Ordenada al origen (2) Pendiente (3) Cuadrado medio residual

25	0.401	1.215
26	0.467	1.415
27	0.316	0.958
28	0.657	1.991
29	0.485	1.470
30	0.494	1.497
31	0.633	1.918
32	0.326	0.988
33	0.651	1.973
34	0.368	1.115
35	0.450	1.364
36	0.563	1.706
37	0.268	0.812
38	0.383	1.161
39	0.355	1.076

3.4. Procedimiento de simulación

A partir del vector de medias y la matriz de covariancia se simuló una muestra de 200 observaciones de una distribución normal multivariada empleando un macro desarrollado para SAS por Lingling (2006) que se presenta en **A2**. En la muestra obtenida, los valores de las variables serológicas generados menores a 0.30 fueron considerados con ese valor. Del mismo modo, los valores de puntajes mayores a 0 fueron considerados 0. Se simularon 15 muestras extraídas de la misma población multivariada empleando distintos números aleatorios como semillas. La base de datos final fue analizada según el modelo propuesto empleando los programas de **A3**.

3.5. Estrategia de análisis

El análisis que se sigue en este trabajo corresponde a un procedimiento en dos etapas recomendado por Anderson y Gerbing (1988). En este enfoque se comienza desarrollando un modelo de medida que ajuste aceptablemente los datos. En esta etapa, se permite que cada variable latente covaríe con cada una de las otras variables latentes, es decir que las tres variables latentes son tratadas como exógenas, covariando entre ellas. El análisis en esta etapa es, esencialmente, un análisis factorial confirmatorio. Si el modelo de medida inicial es inadecuado, se reasignan variables o se eliminan para lograr un mejor ajuste. Una vez que se ha desarrollado el modelo de medida aceptable, se continúa con la segunda etapa en la que se prueba el modelo teórico en sí mismo. En esta etapa se fijan en 0 las covariancias entre las variables latentes que, en este caso, corresponden a la Protección a la Enfermedad y a la Protección a la infección y se establece el camino causal de la variable Protección a la Infección hacia Protección a la Enfermedad que va a ser probado. Si el ajuste del modelo teórico es bueno se comprueban el modelo de medida desarrollado en la primera etapa y, al mismo tiempo, el modelo estructural que es de principal interés. Si el modelo teórico no es satisfactorio, puede ser modificado para obtener un mejor ajuste.

El número de variables manifiestas endógenas (p) es 5 y el de exógenas (q) es 8 y el número de elementos libres y no restringidos en θ (t) en el modelo planteado es 31, por lo tanto

$$t < (1/2)(p + q)(p + q + 1) \text{ en este modelo es } 31 < 91$$

Se cumple la condición necesaria de identificación.

Para verificar condiciones suficientes de identificación se siguen, en este trabajo, procedimientos empíricos, repitiendo el análisis varias veces empleando distintos valores iniciales de los estimadores de los parámetros y verificando que los estimadores finales coincidan.

El método de estimación de los parámetros del modelo empleado es el de Máxima verosimilitud, calculando las estimaciones iniciales mediante la técnica de Mínimos Cuadrados en dos etapas y variables instrumentales.

El criterio de convergencia considerado es el dado por el sistema SAS por defecto, esto es $r = 10^{-\log_{10}(\varepsilon)}$ donde ε está dado por la precisión de la máquina.

El criterio para moverse de un valor al siguiente en la secuencia, esto es, de $\theta^{(i)}$ a $\theta^{(i+1)}$ involucra la función gradiente y la inversa de la matriz hessiana como se detalló en el Capítulo 2.

3.6. Introducción de datos

En este trabajo se empleó la base de datos completa en el programa. El sistema SAS permite agregar la opción COVARIANCE en el procedimiento CALIS para que sea la matriz de covariancia la que se emplee en los cálculos. Si bien la prueba de χ^2 , generalmente, es igual, ya sea que se analice la matriz de correlación o de covariancia, el análisis hecho a partir de la matriz de correlación puede producir errores estándar de los errores no válidos dando pruebas de significación inadecuadas de los coeficientes y otros parámetros (Hatcher, L. 1994).

Capítulo 4: RESULTADOS

4.1. Modelo de medida

Siguiendo la estrategia de dos pasos mencionada en el Capítulo 3, se comenzó evaluando el modelo de medida.

Como se partió de la simulación de una distribución Normal multivariada, no se incluye en este trabajo el análisis de la curtosis y asimetría que confirmó las características de la distribución de los datos.

4.1.1. Modelo de medida sin modificaciones

En una primera etapa se probó el modelo inicial con todas las variables manifiestas y considerando las tres variables latentes covariando entre sí.

Las primeras páginas de la salida del programa SAS incluyen la lista de las variables endógenas y exógenas y la forma general de las ecuaciones estructurales definidas en las sentencias. Estas salidas no se presentan en este trabajo. Se verificó que estos resultados coincidieran con el modelo propuesto. Las siguientes páginas de la salida contienen información descriptiva de las variables manifiestas y la matriz de covariancia a ser analizada. En las páginas que siguen se presentan el vector de los valores iniciales de los estimadores de parámetros como así también la historia de las iteraciones que finalizan en la matriz de covariancia predicha y las matrices de residuales, sin normalizar y normalizados, y también un gráfico de la distribución de los residuales normalizados. Para algunas de las estimaciones iniciales se empleó el método de variables instrumentales.

Como se mencionó en el Capítulo 2, las medidas de ajuste pueden ser globales o individuales. En la Figura 4.1 se reproducen los resultados de la prueba de χ^2 y los índices de ajuste global proporcionados por el programa.

Figura 4.1: Índices de ajuste global del modelo de medida inicial

Covariance Structure Analysis: Maximum Likelihood Estimation	
Fit Function	6.5111
Goodness of Fit Index (GFI)	0.5899
GFI Adjusted for Degrees of Freedom (AGFI)	0.3981
Root Mean Square Residual (RMR)	1.3344
Parsimonious GFI (Mulaik, 1989)	0.4689
Chi-Square	1295.7064
Chi-Square DF	62
Pr > Chi-Square	<.0001
Independence Model Chi-Square	3499.1
Independence Model Chi-Square DF	78
RMSEA Estimate	0.3162

RMSEA 90% Lower Confidence Limit	0.3013
RMSEA 90% Upper Confidence Limit	0.3313
ECVI Estimate	6.8246
ECVI 90% Lower Confidence Limit	6.2354
ECVI 90% Upper Confidence Limit	7.4540
Probability of Close Fit	0.0000
Bentler's Comparative Fit Index	0.6394
Normal Theory Reweighted LS Chi-Square	899.2982
Akaike's Information Criterion	1171.7064
Bozdogan's (1987) CAIC	905.2107
Schwarz's Bayesian Criterion	967.2107
McDonald's (1989) Centrality	0.0458
Bentler & Bonett's (1980) Non-normed Index	0.5463
Bentler & Bonett's (1980) NFI	0.6297
James, Mulaik, & Brett (1982) Parsimonious NFI	0.5005
Z-Test of Wilson & Hilferty (1931)	29.3653
Bollen (1986) Normed Index Rho1	0.5341
Bollen (1988) Non-normed Index Delta2	0.6411
Hoelter's (1983) Critical N	14

Cuando los supuestos apropiados se cumplen (esto es, muestra grande y distribución normal multivariada) la prueba χ^2 constituye una prueba estadística de que el modelo ajusta los datos y por lo tanto se espera un valor- p asociado alto. De acuerdo a los resultados obtenidos el valor del estadístico χ^2 es muy alto (1295.71, el cual, dividido por los grados de libertad, es 20.9) con un valor- p muy pequeño indicando un ajuste global pobre. Como se indicó en el Capítulo 2 existen algunas consideraciones que se deben tener en cuenta a la hora de analizar los resultados de la prueba de χ^2 . Entre estas consideraciones están las referidas al alejamiento de la distribución Normal multivariada y al tamaño de muestra. Asimismo, en esta prueba se asume que el ajuste entre la matriz de covariancia poblacional y la matriz en función de los parámetros del modelo es perfecto. Por esta razón, se complementó el análisis del estadístico χ^2 con los otros índices de ajuste. Como puede observarse en la Figura 4.1, estos índices indican, también, un ajuste pobre con valores bastante menores a 0.90. Las medidas de ajuste global están indicando la necesidad de considerar alguna modificación en el modelo. El análisis de los residuales normalizados indica que la matriz de covariancia predicha por el modelo se aleja más de lo esperable por azar de la matriz de covariancia observada (Figura 4.2).

Figura 4.2: Resultado de la salida del programa SAS con los 10 valores de residuales estandarizados absolutos más altos

Rank Order of the 10 Largest Asymptotically Standardized Residuals

Fila	Columna	Residual
X7	X5	10.38095
Y2	X2	8.22826
Y3	X1	-8.02330
Y5	X1	-7.30049
Y2	X1	-7.10615
X5	X1	7.01706
Y4	X6	6.90014
X5	X4	6.78354
Y4	X8	6.46411
X7	X4	6.44250

Antes de analizar la modificación del modelo, se observaron los resultados del ajuste de los componentes del modelo en forma individual. Como primer paso, se

comprobó que el signo y la magnitud de los coeficientes estimados tuviesen un sentido lógico dentro del contexto del problema, así como también se verificó que no hubiese ningún error estándar de los coeficientes cercanos a cero. En la Figura 4.3 se presenta la salida del programa SAS con los coeficientes y el estadístico t de Student correspondiente a la hipótesis de nulidad de los coeficientes.

Figura 4.3: Coeficientes de las variables observadas y estadístico t

Manifest Variable Equations with Estimates				
X1	=	0.8632*F1	+	1.0000 E1
Std Err		0.0673 Lx1F1		
t Value		12.8261		
X2	=	0.2994*F1	+	1.0000 E2
Std Err		0.0577 Lx2F1		
t Value		5.1896		
X3	=	1.3577*F1	+	1.0000 E3
Std Err		0.1027 Lx3F1		
t Value		13.2221		
X4	=	11.7463*F1	+	1.0000 E4
Std Err		0.5684 Lx4F1		
t Value		20.6653		
X5	=	0.8183*F3	+	1.0000 E5
Std Err		0.0602 Lx5F3		
t Value		13.5915		
X6	=	0.6529*F3	+	1.0000 E6
Std Err		0.0335 Lx6F3		
t Value		19.5156		
X7	=	0.7036*F3	+	1.0000 E7
Std Err		0.0577 Lx7F3		
t Value		12.1864		
X8	=	0.7270*F3	+	1.0000 E8
Std Err		0.0365 Lx8F3		
t Value		19.9248		
Y1	=	0.3516*F2	+	1.0000 Ey1
Std Err		0.0469 Ly1F2		
t Value		7.5026		
Y2	=	1.1699*F2	+	1.0000 Ey2
Std Err		0.1950 Ly2F2		
t Value		5.9989		
Y3	=	2.7851*F2	+	1.0000 Ey3
Std Err		0.2133 Ly3F2		
t Value		13.0552		
Y4	=	2.0572*F2	+	1.0000 Ey4
Std Err		0.2393 Ly4F2		
t Value		8.5960		
Y5	=	9.2996*F2	+	1.0000 Ey5
Std Err		0.4633 Ly5F2		
t Value		20.0745		

Se puede observar que todas las variables son explicadas por las variables latentes que representan como lo indican los valores de los estadísticos t mayores a 2. En la Figura 4.4 se presentan los R^2 que indican la fiabilidad de cada una de las variables indicadoras. Se observa, en estos resultados, que dos de las variancias de errores son negativas. Antes de continuar con el análisis de este modelo, se observan los índices de modificación considerando la posibilidad de modificar el modelo para mejorar el ajuste global. En este sentido, se consideraron la prueba de Wald y los multiplicadores de Lagrange.

Figura 4.4: Fiabilidad de las variables indicadoras en el modelo inicial

	Variable	Varianza del error	Varianza total	R-cuadrado
1	X1	0.52974	1.27482	0.5845
2	X2	0.65784	0.74750	0.1200
3	X3	1.17235	3.01560	0.6112
4	X4	-7.05223	130.92257	1.0539
5	X5	0.38572	1.05527	0.6345
6	X6	0.00911	0.43533	0.9791
7	X7	0.41520	0.91019	0.5438
8	X8	0.0000708	0.52853	0.9999
9	Y1	0.38917	0.51278	0.2411
10	Y2	7.25641	8.62499	0.1587
11	Y3	4.83579	12.59243	0.6160
12	Y4	9.48504	13.71701	0.3085
13	Y5	-5.32395	81.15918	1.0656

Correlaciones entre variables exógenas

Var1	Var2	Parámetro	Estimador
F1	F3	CF1F3	0.72934
F1	F2	CF1F2	0.49921
F3	F2	CF2F3	0.18428

En primer lugar, la prueba de Wald estima el cambio en el χ^2 del modelo que resulta de fijar un parámetro en cero. La Figura 4.5 reproduce la prueba de Wald en la salida del programa SAS.

Figura 4.5: Prueba de Wald para el modelo inicial

Test de Wald multivariante escalonado					
Parámetro	-----Estadísticas acumuladas-----			-Incremento de univarianza-	
	Chi- cuadrado	DF	Pr > ChiSq	Chi- cuadrado	Pr > ChiSq
VARE8	3.11915	1	0.0774	3.11915	0.0774

Los multiplicadores de Lagrange estiman la reducción del estadístico χ^2 del modelo que resulta de permitir que un parámetro sea estimado quitando la restricción impuesta en el modelo. SAS provee dos tipos de matrices de estos multiplicadores; la matriz *Phi* y la matriz *Gamma*. La primera matriz contiene índices para todas las combinaciones posibles de variables latentes y de residuales. Por otra parte, los índices correspondientes a la matriz *Gamma* indican la conveniencia de agregar un nuevo sendero desde una variable latente a alguna indicadora.

La Figura 4.6 contiene los valores de los multiplicadores de Lagrange en orden descendente de importancia. Los resultados en este Cuadro sugieren que el ajuste puede mejorar si se estima la covarianza entre los errores asociados con las variables X6 y X8 (títulos de Ac por SN en bovinos y cobayos, respectivamente) en lugar de considerarla nula.

Figura 4.6: Multiplicadores de Lagrange en orden descendente de importancia

Rank Order of the 10 Largest Lagrange Multipliers in _PHI_

Fila	Columna	Chi-cuadrado	Pr > ChiSq
E8	E6	206.84410	<.0001
E7	E5	107.76088	<.0001
Ey2	E2	73.79972	<.0001
E7	E4	59.29377	<.0001
Ey2	E1	56.23295	<.0001
E7	F3	55.42182	<.0001
E5	E4	54.56303	<.0001
E7	F1	53.98914	<.0001
E5	F3	50.68825	<.0001
E5	F1	46.89712	<.0001

Rank Order of the 10 Largest Lagrange Multipliers in _GAMMA_

Fila	Columna	Chi-cuadrado	Pr > ChiSq
X5	F1	49.60281	<.0001
X1	F2	49.19876	<.0001
X7	F1	47.35182	<.0001
Y4	F3	41.83059	<.0001
X1	F3	32.23262	<.0001
X4	F3	27.45335	<.0001
Y4	F1	19.50220	<.0001
X8	F2	15.50478	<.0001
X6	F2	12.70707	0.0004
X8	F1	10.00466	0.0016

La modificación del modelo agregando senderos desde las variables latentes a las indicadoras de acuerdo al resultado de los multiplicadores de Lagrange de la matriz Gamma no parece una buena alternativa. Por un lado, agregar simplemente el sendero transforma a las variables indicadoras involucradas en variables complejas, afectadas por más de un factor. Por otro lado, cambiar el sendero de una variable latente a otra tampoco es conveniente ya que no existe un sustento teórico para el cambio y, además, todas las variables indicadoras fueron significativas en su relación con el factor correspondiente al modelo.

4.1.2. Modificaciones del modelo de medida

El modelo de medida se rescribió quitando la restricción de la covarianza igual a 0 de los errores de las medidas por seroneutralización. La Figura 4.7 contiene los nuevos valores de los índices para la evaluación del ajuste global del nuevo modelo.

Figura 4.7: Índices de ajuste global del modelo de medida con correlación de errores

Fit Function	5.4061
Goodness of Fit Index (GFI)	0.6284
GFI Adjusted for Degrees of Freedom (AGFI)	0.4457
Root Mean Square Residual (RMR)	1.2655
Parsimonious GFI (Mulaik, 1989)	0.4914
Chi-Square	1075.8110

Chi-Square DF	61
Pr > Chi-Square	<.0001
Independence Model Chi-Square	3499.1
Independence Model Chi-Square DF	78
RMSEA Estimate	0.2891
RMSEA 90% Lower Confidence Limit	0.2741
RMSEA 90% Upper Confidence Limit	0.3044
ECVI Estimate	5.7304
ECVI 90% Lower Confidence Limit	5.1967
ECVI 90% Upper Confidence Limit	6.3044
Probability of Close Fit	0.0000
Bentler's Comparative Fit Index	0.7034
Normal Theory Reweighted LS Chi-Square	764.8793
Akaike's Information Criterion	953.8110
Bozdogan's (1987) CAIC	691.6136
Schwarz's Bayesian Criterion	752.6136
McDonald's (1989) Centrality	0.0791
Bentler & Bonett's (1980) Non-normed Index	0.6207
Bentler & Bonett's (1980) NFI	0.6925
James, Mulaik, & Brett (1982) Parsimonious NFI	0.5416
Z-Test of Wilson & Hilferty (1931)	26.6184
Bollen (1986) Normed Index Rho1	0.6069
Bollen (1988) Non-normed Index Delta2	0.7048
Hoelter's (1983) Critical N	16

En el modelo que incluye la correlación de errores, la diferencia en el valor del estadístico χ^2 con respecto al modelo que no incluye esa correlación es de 219.9 y 1 grado de libertad de diferencia. Este resultado indica que la eliminación de la restricción de la covarianza de los errores de las variables **X6** y **X8** igual a 0, permitiendo estimarla en el modelo, produjo una mejora significativa en el ajuste del modelo. En este mismo sentido, se puede observar una disminución de los valores de los índices de Akaike, Bozdogan y Schwarz y un aumento de los otros índices de ajuste. Los valores de fiabilidad de la Figura 4.8 muestran un cambio en las variables **X6** a **X8**, pero valores similares en las otras variables. Se observa, asimismo, que la estimación de la correlación de los errores es alta.

Figura 4.8: Fiabilidad de las variables indicadoras en el modelo con correlación de errores

Correlaciones múltiples cuadradas				
	Variable	Varianza del error	Varianza total	R-cuadrado
1	X1	0.56647	1.27482	0.5556
2	X2	0.66874	0.74750	0.1054
3	X3	1.25554	3.01560	0.5837
4	X4	-14.53939	130.92257	1.1111
5	X5	0.08518	1.05527	0.9193
6	X6	0.14874	0.43533	0.6583
7	X7	0.14239	0.91019	0.8436
8	X8	0.16311	0.52853	0.6914
9	Y1	0.38608	0.51278	0.2471
10	Y2	7.21582	8.62499	0.1634
11	Y3	4.66735	12.59243	0.6294
12	Y4	9.44175	13.71701	0.3117
13	Y5	-3.48465	81.15918	1.0429

Correlaciones entre variables exógenas

Var1	Var2	Parámetro	Estimador
F1	F3	CF1F3	0.76745
F1	F2	CF1F2	0.50852
F3	F2	CF2F3	0.24088
E6	E8	CE8E9	0.96934

Si bien se ha logrado una mejora en el ajuste con el agregado de la covariancia entre los errores de las variables **X6** y **X8**, el ajuste global sigue siendo pobre. Los errores de las variables **X2** (tiempo de excreción de virus desde la descarga) e **Y2** (tiempo hasta pico de severidad) aparecen en 6 de los 10 multiplicadores mayores de la Figura 4.9, sugiriendo que la eliminación de la restricción en esas covariancias aportaría una mejora en el ajuste. Como no se consideró conveniente agregar más correlaciones de errores para evitar el problema de subidentificación de los parámetros del modelo, se decidió, en primer lugar, eliminar la variable **Y2**, que, si bien fue significativa para representar la variable latente, su valor de fiabilidad fue bajo, y desde el punto de vista teórico no tiene gran relevancia estando las otras indicadores de la protección a la enfermedad en el modelo, especialmente el área bajo la curva de severidad, que expresa el tiempo y severidad de la enfermedad.

Figura 4.9: Multiplicadores de Lagrange en orden descendente de importancia para el modelo con correlación de errores

Rank Order of the 10 Largest Lagrange Multipliers in _PHI_

Fila	Columna	Chi-cuadrado	Pr > ChiSq
Ey2	E2	65.87625	<.0001
Ey2	E1	61.75953	<.0001
Ey2	E5	39.19090	<.0001
Ey4	E2	37.54137	<.0001
Ey4	Ey2	36.84127	<.0001
Ey3	E1	34.93337	<.0001
Ey2	F3	30.28827	<.0001
E1	F2	30.26583	<.0001
E8	E1	28.70265	<.0001
E7	E1	27.54788	<.0001

Al excluir a la variable **Y2** del modelo, el valor de χ^2 fue de 651.27 con 50 grados de libertad, lo que representó una disminución en los valores de los estadísticos χ^2 de 424.54 con una diferencia de 11 grados de libertad con respecto al modelo anterior. Esta disminución fue significativa, indicando una mejora con la modificación del modelo.

En la etapa siguiente se modificó el modelo eliminando a la variable **X2** por las mismas razones por las que se realizó la modificación anterior del modelo. El valor del estadístico χ^2 fue de 465.49 con 40 grados de libertad. La eliminación de la variable **X2** produjo una mejora significativa en el modelo. En el Cuadro 4.1 se presentan las distintas medidas de evaluación global. Se pueden observar como fueron mejorando los distintos índices con las sucesivas modificaciones del modelo. Si bien aún no se llegaron a obtener valores satisfactorios de ajuste, la implementación de nuevos cambios no podrían justificarse desde el punto de vista de la teoría, como sería agregar nuevas sendas desde variables latentes a indicadoras, o podrían provocar problemas de

identificación del modelo como sería el agregado de correlación entre errores o la eliminación de alguna otra variable indicadora.

Cuadro 4.1: Comparación de modelos de medidas de evaluación global

Medidas de Ajuste Global	Modelo⁽¹⁾			
	1	2	3	4
χ^2	1295.71	1075.81	651.27	465.49
Grados de libertad	62	61	50	40
Akaike	1171.71	953.81	551.27	385.49
Bozdogan	905.21	691.61	336.35	213.56
Schwarz	967.21	752.61	386.35	253.56
Bentler & Bonett's (1980) Non-normed Index	0.55	0.69	0.73	0.79
Bollen (1988) Non-normed Index Delta2	0.64	0.70	0.80	0.85
GFI	0.59	0.63	0.70	0.73
GFI Adjusted for Degrees of Freedom (AGFI)	0.40	0.45	0.54	0.56
Parsimonious GFI (Mulaik, 1989)	0.47	0.49	0.53	0.53
Bentler's Comparative Fit Index	0.64	0.70	0.80	0.85

(1) Modelo 1: modelo inicial Modelo 2: modelo con estimación de covariancia de errores Modelo 3: modelo con estimación de covariancia de errores sin Y2 Modelo 4: modelo con estimación de covariancia de errores sin Y2 y X2.

Un problema que se observó, y que no fue tratado hasta aquí, fue la presencia de dos correlaciones mayores a 1 y variancias de error negativas relacionadas a estos casos. Estos casos son comunes en este tipo de análisis y son conocidos como “casos Heywood” y fueron mencionados en el Capítulo 2. Este problema ocurre con variables que se espera tengan coeficientes altos, por lo que se descartó la eliminación de las variables involucradas como posible solución. Por otra parte, como no hubo problemas de convergencia, ni presencia de valores atípicos se optó por ignorar el valor negativo y considerarlo esencialmente 0, fijando las variancias de esos errores en valores muy bajos (0.0000001).

4.1.3 Modelo de medida final

4.1.3.1 Medidas de ajuste global

La Figura 4.10 reproduce la salida de computadora del programa de SAS con los valores de las medidas de ajuste global del modelo de medida final. Se puede observar

que las medidas de ajuste global sufrieron pequeños cambios con el agregado de la restricción de las variancias de errores.

Figura 4.10: Índices de ajuste global del modelo de medida final

Fit Function	2.5681
Goodness of Fit Index (GFI)	0.6981
GFI Adjusted for Degrees of Freedom (AGFI)	0.5256
Root Mean Square Residual (RMR)	1.3392
Parsimonious GFI (Mulaik, 1989)	0.5331
Chi-Square	511.0574
Chi-Square DF	42
Pr > Chi-Square	<.0001
Independence Model Chi-Square	2827.8
Independence Model Chi-Square DF	55
RMSEA Estimate	0.2369
RMSEA 90% Lower Confidence Limit	0.2187
RMSEA 90% Upper Confidence Limit	0.2555
ECVI Estimate	2.8248
ECVI 90% Lower Confidence Limit	2.4676
ECVI 90% Upper Confidence Limit	3.2220
Probability of Close Fit	0.0000
Bentler's Comparative Fit Index	0.8308
Normal Theory Reweighted LS Chi-Square	473.3661
Akaike's Information Criterion	427.0574
Bozdogan's (1987) CAIC	246.5281
Schwarz's Bayesian Criterion	288.5281
McDonald's (1989) Centrality	0.3095
Bentler & Bonett's (1980) Non-normed Index	0.7785
Bentler & Bonett's (1980) NFI	0.8193
James, Mulaik, & Brett (1982) Parsimonious NFI	0.6256
Z-Test of Wilson & Hilferty (1931)	17.9457
Bollen (1986) Normed Index Rho1	0.7633
Bollen (1988) Non-normed Index Delta2	0.8316
Hoelter's (1983) Critical N	24

La función de ajuste minimizada fue de 2.5681, valor que multiplicado por el número de observaciones menos 1 (199) reproduce el valor del estadístico χ^2 (511.0574). Este valor lleva al rechazo de la hipótesis de que la matriz de covariancia poblacional puede ser escrita como función de los parámetros del modelo propuesto.

Los índices que pueden clasificarse como absolutos como GFI, AGFI, el índice de Holter, ECVI, RMR, tienen problemas similares a los de χ^2 porque se basan en simples variaciones de χ^2 .

Los índices de tipo relativo comparan el χ^2 del modelo probado con uno del llamado “modelo nulo” (también llamado “modelo de independencia”). El modelo nulo es el modelo que especifica que las variables no están correlacionadas entre sí (no hay variables latentes). El χ^2 correspondiente a este modelo es 2827.8, es un valor muy grande que indica un ajuste muy pobre del modelo de independencia. Entre estos índices que indican la mejora en el ajuste al considerar el modelo propuesto en lugar del nulo, se encuentra el NFI. De acuerdo a los resultados de la Figura 4.10, se observa que se obtiene una reducción del 82% de la función de ajuste o del χ^2 cuando se pasa del modelo nulo al modelo propuesto, es decir que se obtiene un 82% de mejora en el ajuste cuando se considera el modelo propuesto. Con una interpretación similar, el índice de Bollen (1988) NNI delta 2, menos dependiente de los grados de libertad tuvo un valor de 0.83. Otro índice para evaluar el nivel de mejora en el ajuste con relación al modelo nulo es el Rho1, propuesto por Bollen (1986), que permite comparar modelos con los

misimos valores en la función de ajuste pero especificaciones más complejas, favoreciendo a aquellos más simples.

Ajuste del modelo por variable latente

Una estrategia propuesta por Bollen (1989) para la re-especificación del modelo es lo que él denomina “ajuste del modelo por partes”. Este procedimiento se siguió en este trabajo y consistió en particionar el modelo de ajuste pobre en las tres variables latentes re-estimando cada parte por separado de modo de identificar la componente del modelo que no ajustaba adecuadamente. En una etapa posterior se analizaron las variables latentes de a dos. En el Cuadro 4.2 se presentan algunos de los índices de ajuste global para cada variable latente por separado.

Cuadro 4.2: Medidas de ajuste global para cada componente del modelo

Medidas de Ajuste Global	Variable latente		
	Protección a la infección ¹	Protección a la enfermedad ²	Calidad por laboratorio ³
χ^2	35.95	40.17	16.36
Grados de libertad	2	2	1
Bentler & Bonett's (1980) Non-normed Index	0.92	0.71	0.93
Bollen (1988) Non-normed Index Delta2	0.93	0.90	0.99
GFI	0.92	0.91	0.96
Bentler's Comparative Fit Index	0.93	0.90	0.99

1 Se incluyeron todas las variables originales; 2 Sin incluir Y2; 3 Incluyendo la covariancia de errores de las variables de SN

Cuando se analizó el ajuste del modelo por partes, considerando individualmente cada variable latente, los índices de ajuste fueron superiores a 0.90, con excepción de el índice de Bentler & Bonett (1980) para la variable Protección a la Enfermedad que alcanzó un valor muy bajo. Los índices alcanzaron valores muy altos para la variable latente de Calidad por Laboratorio. El peor ajuste correspondió a la variable Protección a la Enfermedad.

En la siguiente etapa, se analizaron las variables latentes de a pares, obteniéndose los resultados que se resumen en el Cuadro 4.3.

Cuadro 4.3: Medidas de ajuste global para los modelos con las variables latentes de a pares

Medidas de Ajuste Global	Variables latentes ¹		
	Protección a la infección y Calidad por laboratorio	Protección a la infección y Protección a la enfermedad	Protección a la enfermedad y Calidad por laboratorio
χ^2	176.65	184.38	179.59
Grados de libertad	12	13	18
Bentler & Bonett's (1980) Non-normed Index	0.91	0.82	0.90
Bollen (1988) Non-normed Index Delta2	0.92	0.83	0.91
GFI	0.81	0.79	0.83
Bentler's Comparative Fit Index	0.92	0.83	0.91
Correlación	0.78 *	0.46 *	0.19 *

¹ Se consideró el modelo final

Se observan problemas en el ajuste cuando se consideran las variables Protección a la Infección y Protección a la Enfermedad, con valores de índices de ajuste inferiores a 0.90, indicando un ajuste pobre. Cuando se considera cada una de estas variables por separado y Calidad por Laboratorio, los índices alcanzan, en general, valores satisfactorios, indicando un ajuste aceptable. Las correlaciones para cada par de variables latentes fueron significativas.

Los valores de los coeficientes estimados de las variables indicadoras fueron similares a los obtenidos con el modelo completo, tanto cuando se consideraron las variables latentes individualmente, como cuando se las consideró de a pares. Estos resultados no se reproducen en este trabajo.

4.2 Medidas de Ajuste de los componentes del modelo

4.2.1 Estimadores de los parámetros

En la Figura 4.11 se reproducen los valores de los coeficientes estimados.

Figura 4.11. Salidas de SAS de las ecuaciones con las variables manifiestas y sus estimaciones en el modelo de medida

a) Valores no estandarizados

```

X1      =  0.8697*F1      +  1.0000 E1
Std Err  0.0671 Lx1F1
t Value  12.9570
X3      =  1.3959*F1      +  1.0000 E3
Std Err  0.1013 Lx3F1
t Value  13.7817
X4      =  11.4421*F1     +  1.0000 E4
Std Err  0.5735 Lx4F1
t Value  19.9499
X5      =  0.9831*F3      +  1.0000 E5
Std Err  0.0543 Lx5F3
t Value  18.0900
X6      =  0.5402*F3      +  1.0000 E6
Std Err  0.0387 Lx6F3
t Value  13.9586
X7      =  0.8719*F3      +  1.0000 E7
Std Err  0.0523 Lx7F3
t Value  16.6807
X8      =  0.6109*F3      +  1.0000 E8
Std Err  0.0420 Lx8F3
t Value  14.5374
Y1      =  0.3631*F2      +  1.0000 Ey1
Std Err  0.0474 Ly1F2
t Value  7.6636
Y3      =  2.8585*F2      +  1.0000 Ey3
Std Err  0.2068 Ly3F2
t Value  13.8254
Y4      =  2.0890*F2      +  1.0000 Ey4
Std Err  0.2408 Ly4F2
t Value  8.6768
Y5      =  9.0088*F2      +  1.0000 Ey5
Std Err  0.4516 Ly5F2
t Value  19.9499

```

Variances of Exogenous Variables

Variable	Parámetro	Estimador	Error estándar	Valor t
F1		1.00000		
F3		1.00000		
F2		1.00000		
E1	VARE1	0.51840	0.05197	9.97
E3	VARE3	1.06720	0.10699	9.97
E4		1E-7		
E5	VARE5	0.08886	0.01956	4.54
E6	VARE6	0.14354	0.01604	8.95
E7	VARE7	0.15002	0.02082	7.21
E8	VARE8	0.15533	0.01771	8.77
Ey1	VAREy1	0.38091	0.03819	9.97
Ey3	VAREy3	4.42140	0.44325	9.97
Ey4	VAREy4	9.35301	0.93765	9.97
Ey5		1E-7		

Covariances Among Exogenous Variables

Var1	Var2	Parámetro	Estimador	Error estándar	Valor t
F1	F3	CF1F3	0.83637	0.02342	35.71
F1	F2	CF1F2	0.52611	0.05127	10.26
F3	F2	CF2F3	0.25638	0.06813	3.76
E6	E8	CE8E9	0.14460	0.01662	8.70

b) Valores estandarizados

X1	=	0.7703 *F1	+	0.6377	E1
		Lx1F1			
X3	=	0.8038 *F1	+	0.5949	E3
		Lx3F1			
X4	=	1.0000 *F1	+	0.000028	E4
		Lx4F1			
X5	=	0.9570 *F3	+	0.2902	E5
		Lx5F3			
X6	=	0.8187 *F3	+	0.5742	E6
		Lx6F3			
X7	=	0.9139 *F3	+	0.4060	E7
		Lx7F3			
X8	=	0.8403 *F3	+	0.5421	E8
		Lx8F3			
Y1	=	0.5071 *F2	+	0.8619	Ey1
		Ly1F2			
Y3	=	0.8055 *F2	+	0.5925	Ey3
		Ly3F2			
Y4	=	0.5640 *F2	+	0.8257	Ey4
		Ly4F2			
Y5	=	1.0000 *F2	+	0.000035	Ey5
		Ly5F2			

Correlaciones entre variables exógenas

Var1	Var2	Parámetro	Estimador
F1	F3	CF1F3	0.83637
F1	F2	CF1F2	0.52611
F3	F2	CF2F3	0.25638
E6	E8	CE8E9	0.96842

Los coeficientes indican el efecto directo de la variable latente en las variables observadas. En todos los casos se observan valores de t de Student mayores a 2, indicando que todos los efectos son significativos, es decir que las variables latentes explican la variabilidad de las variables observadas. Los coeficientes estandarizados tienen valores en el rango de 0.50 a 1.00, y con la excepción de **Y1** (pico de severidad de la rinitis) e **Y4** (severidad a los 14 días post-descarga), todos los coeficientes son superiores a 0.60, es decir que son moderadamente altos a muy altos. Se destacan los coeficientes de las variables de **área bajo la curva**, tanto de excreción de virus (**X4**), como de síntomas clínicos (**Y5**). El resto de los coeficientes asociados con la variable de protección a la infección son altos. Asimismo, son altos todos los coeficientes de las variables asociadas a la medición de anticuerpos en laboratorio, siendo superiores los obtenidos cuando se empleó la técnica de ELISA en lugar de la técnica de seroneutralización. Los coeficientes correspondientes a las variables pico de severidad y severidad a los 14 días (**Y1** e **Y4**, respectivamente) presentaron los coeficientes estandarizados más bajos.

El coeficiente de correlación entre las variables protección a la infección y calidad por el laboratorio fue de 0.84, indicando una fuerte relación entre estas variables latentes.

Aunque significativo, el coeficiente de correlación entre las variables latentes Protección a la Infección y Protección a la Enfermedad fue moderadamente bajo (0.53), y fue aún más bajo el coeficiente de correlación entre las variables Protección a la Enfermedad y Calidad por el Laboratorio (0.26).

La correlación entre los errores de las variables que evalúan anticuerpos por seroneutralización fue muy fuerte (0.97).

4.2.2 Otras medidas de evaluación de los componentes del modelo de medida

En el Cuadro 4.4 se presentan los valores de las medidas de fiabilidad (R^2) para cada variable observada, las medidas de fiabilidad compuesta y de la variancia extraída para cada variable latente.

Las medidas de fiabilidad de las variables manifiestas indican el porcentaje de la variación de la variable que es explicada por el factor que supone que mide. Los valores de fiabilidad fueron de 0.26 a 1. El factor que presentó valores más bajos de fiabilidad de las variables indicadoras fue la Protección a la Enfermedad, los otros dos factores presentaron valores más altos de fiabilidades individuales.

El índice de fiabilidad compuesta refleja la consistencia interna de los indicadores que miden el factor. Los valores obtenidos fueron todos superiores a 0.80, superando ampliamente los límites mínimos requeridos para una fiabilidad aceptable.

Los índices de variancia extraída, que miden la cantidad de variabilidad que es capturada por el factor en relación a la variancia del error de medida, superaron, en todos los casos, el valor mínimo sugerido de 0.50. El factor Protección a la Enfermedad fue el que presentó el valor más bajo de variancia extraída.

Cuadro 4.4: Propiedades del modelo de medida

Variable latente y Variables indicadoras	Coefficiente estandarizado	t^a	Fiabilidad	Variancia extraída
F1 Protección a la infección			0.897 ^b	0.746
X1 Pico excretorio	0.770	12.96	0.593	
X3 Duración de excreción del virus	0.804	13.78	0.646	
X4 Área bajo la curva de excreción de virus	1.000	19.95	1.000	
F2 Protección a la enfermedad			0.823	0.556
Y1 Pico de severidad	0.507	7.66	0.257	
Y3 Duración de síntomas	0.806	13.83	0.649	
Y4 Puntaje a los 14 días	0.564	8.68	0.318	
Y5 Área bajo la curva de severidad	1.000	19.95	1.000	
F3 Calidad por laboratorio			0.934	0.782

X5 Título de Ac por ELISA en bovino	0.957	18.09	0.916	
X6 Título de Ac por SN en bovino	0.819	13.96	0.670	
X7 Título de Ac por ELISA en cobayo	0.914	16.68	0.835	
X8 Título de Ac por SN en cobayo	0.840	14.54	0.706	

^a todas las pruebas t fueron significativas al nivel de significación < 0.001

^b Indica fiabilidad compuesta

4.3 Modelo estructural

Siguiendo con la estrategia de análisis de dos pasos y, habiendo analizado y seleccionado las variables indicadoras para medir cada variable latente, se continuó con el segundo paso, modificando el modelo de modo de analizar la estructura causal de los factores Protección a la Infección y Protección a la Enfermedad. El modelo teórico establece que la enfermedad es causada por la infección, es decir que si se protege contra la infección se espera protección contra la enfermedad por lo que, en un diagrama de senderos, una flecha recta de una sola punta iría desde esta variable latente hacia la primera. Esto significa que la variable latente Protección a la enfermedad pasa a ser, ahora, una variable latente endógena y, por lo tanto, requiere que se la incluya en una ecuación. En el modelo teórico las variables latentes Protección a la infección y Calidad por el laboratorio covarían, siendo ambas consideradas exógenas.

Otra cambio en el modelo que se implementó se refiere a la indeterminación de la escala de las variables latentes. Como las variables latentes son constructos hipotéticos, no tienen una métrica o escala definida, resulta imposible distinguir entre situaciones en las que el factor tiene variancia grande y los coeficientes para ese factor son pequeños o viceversa. Para resolver este problema, se le dio a todos los factores, variancias unitarias, es decir que se fijaron en 1 las variancias de los factores en el modelo de medida. De este modo, se estableció una escala para los factores y se pudo asegurar que el modelo pudiese ser identificado. En el modelo estructural se requiere que las variancias de los factores exógenos sean parámetros que se estimen libremente por lo cual se hace necesario abordar nuevamente el problema de la indeterminación de escala. En este sentido, se debe fijar en 1 los coeficientes de la variable indicadora que, en cierto sentido, mejor representa al constructo latente (Joreskog y Sorbom, 1989). En este trabajo se decidió fijar en 1 los coeficientes que habían dado valores más altos en el análisis del modelo de medida. Esta decisión estaba en correspondencia con la importancia teórica de las variables observadas en representar el factor. De este modo, se fijaron en 1 los coeficientes de las variables área bajo la curva de excreción de virus y de síntomas (que habían dado 1 ya en el análisis del modelo de medida) y el coeficiente para la variable anticuerpos por ELISA en bovinos.

En el modelo estructural, al considerar a la variable latente Protección a la enfermedad como endógena e incluirla en una ecuación, se debió estimar la variancia

del error para esa variable. Asimismo, las variancias de las variables latentes exógenas no se fijaron si no que se estimaron libremente.

4.3.1 Ajuste global del modelo estructural

Del mismo modo que en la primera etapa del análisis, la Figura 4.12 presenta la salida del *software* con el resumen de las medidas de evaluación global del modelo. Las medidas de ajuste global fueron muy similares a las obtenidas con el modelo de medida. El ajuste sigue siendo pobre.

Figura 4.12: Salida del software con el resumen de las medidas de ajuste global del modelo estructural

Covariance Structure Analysis: Maximum Likelihood Estimation

Fit Function	2.7059
Goodness of Fit Index (GFI)	0.6790
GFI Adjusted for Degrees of Freedom (AGFI)	0.5073
Root Mean Square Residual (RMR)	1.3651
Parsimonious GFI (Mulaik, 1989)	0.5308
Chi-Square	538.4720
Chi-Square DF	43
Pr > Chi-Square	<.0001
Independence Model Chi-Square	2827.8
Independence Model Chi-Square DF	55
RMSEA Estimate	0.2406
RMSEA 90% Lower Confidence Limit	0.2227
RMSEA 90% Upper Confidence Limit	0.2590
ECVI Estimate	2.9519
ECVI 90% Lower Confidence Limit	2.5844
ECVI 90% Upper Confidence Limit	3.3592
Probability of Close Fit	0.0000
Bentler's Comparative Fit Index	0.8213
Normal Theory Reweighted LS Chi-Square	517.4503
Akaike's Information Criterion	452.4720
Bozdogan's (1987) CAIC	267.6443
Schwarz's Bayesian Criterion	310.6443
McDonald's (1989) Centrality	0.2898
Bentler & Bonett's (1980) Non-normed Index	0.7714
Bentler & Bonett's (1980) NFI	0.8096
James, Mulaik, & Brett (1982) Parsimonious NFI	0.6329
Z-Test of Wilson & Hilferty (1931)	18.4642
Bollen (1986) Normed Index Rho1	0.7564
Bollen (1988) Non-normed Index Delta2	0.8221
Hoelter's (1983) Critical N	23

En la inspección de los residuales normalizados se observa que se encuentran residuales con valores absolutos mayores a 2, resultado que está en concordancia con el pobre ajuste global. En la Figura 4.13 se presenta la salida de software con los valores de los residuales extremos. Se puede observar que 9 de 10 de estos valores involucran las variables manifiestas X1 (pico excretorio) y Y4 (síntomas a los 14 días). Por el modelo propuesto se están sobreestimando las covariancias entre X1 con Y3, con Y5 y con X3, y se están subestimando las covariancias entre X1 con X5, con X6 y con X8. Estas últimas variables son las que representan la Calidad por laboratorio (ELISA y Seroneutralización en bovinos y Seroneutralización en cobayos). Por otra parte, el modelo está subestimando las covariancias entre la variable Y4 con X6, Y1 y X8.

Nuevamente, están involucradas aquí las variables de Calidad por laboratorio medidas por Seroneutralización, tanto en bovinos como en cobayos.

Figura 4.13: Salida del *software* con los 10 residuales normalizados extremos

Fila	Columna	Residual
Y3	X1	-8.81066
Y5	X1	-7.88091
X5	X3	-6.20585
X8	X1	5.43201
Y4	X6	5.09353
Y4	Y1	5.02284
X3	X1	-4.87036
X5	X1	4.63381
Y4	X8	4.60651
X6	X1	4.60349

4.3.2 Ajuste de los componentes del modelo

Los errores estándar estimados de los coeficientes de los factores no tuvieron valores cercanos a 0, que indicarían problemas de estimación, como puede observarse en la Figura 4.14.

Figura 4.14: Salida del *software* con las ecuaciones de las variables manifiestas y latentes endógenas en el modelo estructural

a) Valores no estandarizados

Manifest Variable Equations with Estimates

X1	=	0.0760*F1	+	1.0000	E1
Std Err		0.00446 Lx1F1			
t Value		17.0401			
X3	=	0.1220*F1	+	1.0000	E3
Std Err		0.00640 Lx3F1			
t Value		19.0609			
X4	=	1.0000 F1	+	1.0000	E4
X5	=	1.0000 F3	+	1.0000	E5
X6	=	0.5410*F3	+	1.0000	E6
Std Err		0.0311 Lx6F3			
t Value		17.4159			
X7	=	0.8793*F3	+	1.0000	E7
Std Err		0.0361 Lx7F3			
t Value		24.3722			
X8	=	0.6114*F3	+	1.0000	E8
Std Err		0.0329 Lx8F3			
t Value		18.5646			
Y1	=	0.0403*F2	+	1.0000	Ey1
Std Err		0.00486 Ly1F2			
t Value		8.3005			
Y3	=	0.3173*F2	+	1.0000	Ey3
Std Err		0.0165 Ly3F2			
t Value		19.1772			
Y4	=	0.2319*F2	+	1.0000	Ey4
Std Err		0.0241 Ly4F2			
t Value		9.6359			
Y5	=	1.0000 F2	+	1.0000	Ey5

Latent Variable Equations with Estimates

F2 = 0.4142*F1 + 1.0000 D1
 Std Err 0.0475 PF2F1
 t Value 8.7271

b)Valores estandarizados

Manifest Variable Equations with Standardized Estimates

X1 = 0.7703*F1 + 0.6377 E1
 Lx1F1
 X3 = 0.8038*F1 + 0.5949 E3
 Lx3F1
 X4 = 1.0000 F1 +0.000028 E4
 X5 = 0.9638 F3 + 0.2665 E5
 X6 = 0.8118*F3 + 0.5839 E6
 Lx6F3
 X7 = 0.9125*F3 + 0.4090 E7
 Lx7F3
 X8 = 0.8326*F3 + 0.5538 E8
 Lx8F3
 Y1 = 0.5071*F2 + 0.8619 Ey1
 Ly1F2
 Y3 = 0.8055*F2 + 0.5925 Ey3
 Ly3F2
 Y4 = 0.5640*F2 + 0.8257 Ey4
 Ly4F2
 Y5 = 1.0000 F2 +0.000035 Ey5

Latent Variable Equations with Standardized Estimates

F2 = 0.5261*F1 + 0.8504 D1
 PF2F1

Los resultados son similares a los obtenidos en el modelo de medida. Es de especial interés en el modelo estructural analizar el coeficiente correspondiente al sendero de la variable latente Protección a la infección a la Protección a la enfermedad. Se puede observar en la Figura 4.14a que este coeficiente es significativo a un nivel de significación < 0.001 .

En la Figura 4.15 se reproduce la salida del software con los valores de los R^2 de las variables. Se observa que el porcentaje de variabilidad de la variable latente Protección a la enfermedad explicada por la Protección a la infección es cercana al 30%.

Figura 4.15: Salida del *software* con los valores de R^2 de las variables endógenas y los coeficientes de correlación de las variables exógenas

	Variable	Varianza del error	Varianza total	R-cuadrado
1	X1	0.51840	1.27482	0.5934
2	X3	1.06720	3.01560	0.6461
3	X4	1E-7	130.92257	1.0000
4	X5	0.07493	1.05527	0.9290
5	X6	0.14844	0.43533	0.6590
6	X7	0.15226	0.91019	0.8327
7	X8	0.16212	0.52853	0.6933
8	Y1	0.38091	0.51278	0.2572
9	Y3	4.42140	12.59243	0.6489
10	Y4	9.35301	13.71701	0.3181

11	Y5	1E-7	81.15918	1.0000
12	F2	58.69513	81.15918	0.2768

Correlaciones entre variables exógenas

Var1	Var2	Parámetro	Estimador
F1	F3	CF1F3	0.83221
E6	E8	CE8E9	0.96936

4.4 Ajuste de la componente estructural del modelo

Para la evaluación de la componente estructural del modelo se calculó el índice de ajuste relativo normado (RNFI), estimando un modelo en el que las covariancias entre variables latentes se fijaron en 0, y considerando los resultados obtenidos con el modelo de medida y estructural. En el Cuadro 4.5 se resumen los resultados de los modelos incluyendo al modelo nulo que considera que las variables no están correlacionadas entre sí.

Cuadro 4.5: Resumen del ajuste de modelos

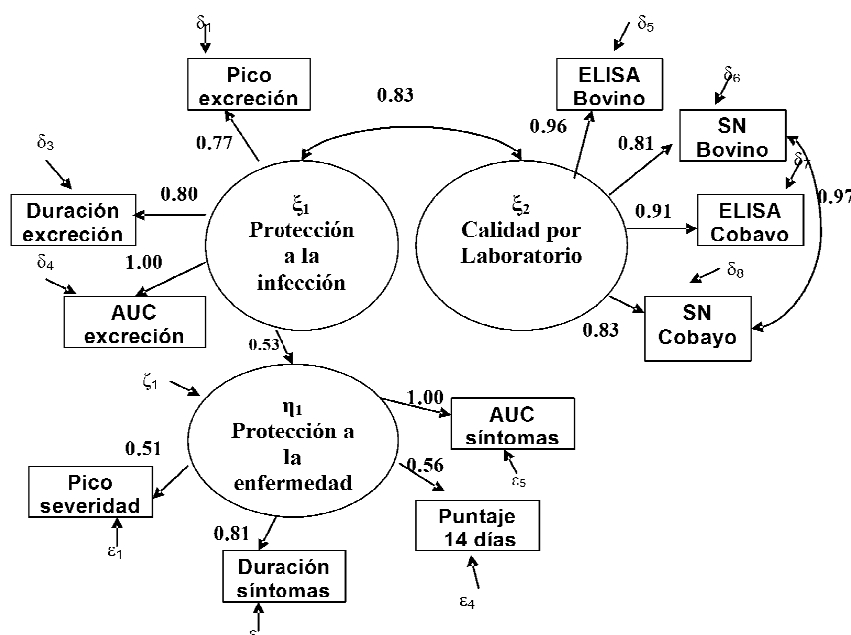
Modelo	Modelo combinado						Modelo Estructural
	χ^2	gl	NFI	NNFI	CFI	PNF1	RNFI
Nulo	2827.8	55	0.00	--	--	--	--
Variables latentes no correlacionadas	813.43	45	0.71	0.66	0.72	0.58	0.00
Modelo estructural	538.47	43	0.81	0.77	0.82	0.63	0.91
Modelo de medida	511.06	42	0.82	0.78	0.83	0.63	1.00

NFI: Índice de ajuste normado; NNFI: Índice de ajuste no normado; CFI: Índice de ajuste comparativo; PNF1: Índice de ajuste normado de parsimonia; RNFI: Índice de ajuste relativo normado.

El valor del RNFI, 0.91, indica un ajuste aceptable del componente estructural del modelo.

La Figura 4.16 representa el diagrama de sendero final con los parámetros estimados.

Figura 4.16: Diagrama de senderos final y coeficientes estandarizados estimados



4.5 Resultado de simulaciones de 15 muestras

En el Cuadro 4.6 se presentan los resultados del ajuste del modelo estructural propuesto obtenido a partir de la selección al azar de 15 muestras extraídas de la población multivariada simulada. En todos los casos, la componente estructural del modelo ajustó adecuadamente. Solo dos muestras presentan un RNFI de 0.88, pero en el resto de las muestras este índice supera el límite de 0.90, considerado como indicador de buen ajuste a los datos.

Cuadro 4.6: Índices de ajuste global del modelo estructural resultados de 15 muestras extraídas aleatoriamente de la población simulada

Simulación	Chi-cuadrado	NFI	NNFI	CFI	RNFI
1	538.47	0.81	0.77	0.82	0.91
2	674.72	0.79	0.74	0.80	0.90
3	501.08	0.83	0.79	0.84	0.92
4	543.09	0.83	0.80	0.84	0.88
5	515.09	0.83	0.81	0.84	0.99
6	640.03	0.79	0.74	0.80	0.88
7	495.30	0.83	0.80	0.84	0.95
8	500.64	0.83	0.80	0.84	0.97
9	531.43	0.82	0.79	0.84	0.93
10	464.64	0.84	0.81	0.85	0.95
11	515.03	0.82	0.78	0.83	0.92

12	642.07	0.77	0.72	0.78	0.95
13	512.67	0.83	0.80	0.85	0.96
14	592.93	0.80	0.76	0.81	0.95
15	565.11	0.82	0.78	0.83	0.93

Los coeficientes estandarizados y la fiabilidad de las variables de las 15 muestras simuladas se presentan en el Cuadro 4:7. Los coeficientes más bajos corresponden a la variable latente Protección a la enfermedad. El coeficiente de correlación entre las variables latentes Protección a la infección y Calidad según el laboratorio tuvo un valor mínimo de 0.79 y un valor máximo de 0.89. El coeficiente para la variable latente Protección a la enfermedad fue relativamente bajo, con un valor mínimo de 0.43 y un valor máximo de 0.60.

Cuadro 4.7: Coeficientes estandarizados y fiabilidades de las variables según el modelo estructural resultado de 15 muestras extraídas de una población simulada

	Simulación														
Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
F1: Protección a la infección															
X1	0.77	0.80	0.79	0.77	0.79	0.77	0.79	0.78	0.80	0.78	0.76	0.74	0.83	0.77	0.76
Pico excretorio	(0.59)	(0.65)	(0.62)	(0.59)	(0.63)	(0.60)	(0.63)	(0.62)	(0.64)	(0.61)	(0.58)	(0.54)	(0.69)	(0.60)	(0.58)
X3	0.80	0.84	0.80	0.85	0.80	0.77	0.80	0.82	0.84	0.79	0.76	0.78	0.82	0.79	0.85
Duración de excreción	(0.65)	(0.70)	(0.65)	(0.72)	(0.64)	(0.59)	(0.64)	(0.67)	(0.70)	(0.62)	(0.58)	(0.61)	(0.67)	(0.62)	(0.72)
X4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AUC excreción	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
F2: Protección a la enfermedad															
Y1	0.51	0.47	0.58	0.55	0.58	0.57	0.47	0.59	0.53	0.47	0.49	0.46	0.48	0.62	0.58
Pico de severidad	(0.26)	(0.22)	(0.34)	(0.30)	(0.34)	(0.32)	(0.22)	(0.34)	(0.29)	(0.22)	(0.24)	(0.21)	(0.23)	(0.39)	(0.33)
Y3	0.81	0.84	0.82	0.86	0.84	0.82	0.78	0.84	0.80	0.78	0.79	0.80	0.78	0.82	0.82
Duración de síntomas	(0.65)	(0.70)	(0.67)	(0.73)	(0.71)	(0.68)	(0.62)	(0.71)	(0.63)	(0.60)	(0.63)	(0.64)	(0.61)	(0.67)	(0.67)
Y4	0.56	0.60	0.65	0.67	0.64	0.61	0.63	0.66	0.66	0.61	0.57	0.61	0.66	0.63	0.71
Puntaje 14 días	(0.32)	(0.37)	(0.42)	(0.45)	(0.42)	(0.37)	(0.40)	(0.43)	(0.43)	(0.38)	(0.33)	(0.37)	(0.43)	(0.39)	(0.51)
Y5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AUC síntomas	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
F3: Calidad por laboratorio															
X5	0.96	0.97	0.96	0.97	0.97	0.95	0.99	0.97	0.98	0.97	0.96	0.97	0.97	0.98	0.97
ELISA bovinos	(0.93)	(0.95)	(0.92)	(0.94)	(0.95)	(0.90)	(0.98)	(0.95)	(0.96)	(0.94)	(0.92)	(0.94)	(0.95)	(0.96)	(0.94)
X6	0.81	0.81	0.81	0.80	0.87	0.85	0.82	0.83	0.79	0.83	0.80	0.79	0.84	0.82	0.90

SN	(0.66)	(0.65)	(0.65)	(0.64)	(0.75)	(0.73)	(0.67)	(0.68)	(0.62)	(0.68)	(0.65)	(0.62)	(0.70)	(0.67)	(0.80)
Bovinos															
X7	0.91	0.94	0.94	0.94	0.94	0.94	0.94	0.93	0.94	0.94	0.94	0.92	0.96	0.93	0.94
ELISA cobayos	(0.83)	(0.89)	(0.88)	(0.87)	(0.89)	(0.89)	(0.89)	(0.86)	(0.88)	(0.89)	(0.88)	(0.85)	(0.92)	(0.87)	(0.88)
X8	0.83	0.82	0.83	0.81	0.88	0.86	0.83	0.84	0.80	0.83	0.82	0.81	0.85	0.82	0.90
SN cobayos	(0.69)	(0.67)	(0.69)	(0.66)	(0.77)	(0.75)	(0.69)	(0.70)	(0.65)	(0.69)	(0.68)	(0.66)	(0.73)	(0.68)	(0.82)
F1→F2	0.53	0.54	0.51	0.62	0.51	0.43	0.46	0.57	0.60	0.53	0.44	0.44	0.45	0.52	0.60
	(0.28)	(0.29)	(0.26)	(0.39)	(0.26)	(0.18)	(0.21)	(0.32)	(0.36)	(0.29)	(0.20)	(0.20)	(0.20)	(0.27)	(0.36)
F1 F3¹	0.83	0.84	0.83	0.83	0.82	0.83	0.80	0.83	0.81	0.85	0.86	0.79	0.89	0.81	0.88

1 Correlación entre variables latentes exógenas

Capítulo 5: DISCUSIÓN

El control de Rinotraqueitis bovina en nuestro país se basa fundamentalmente en el uso de vacunas inactivadas polivalentes de diferentes marcas nacionales e importadas. Para garantizar la efectividad de las vacunas es necesario implementar pruebas que valoren objetivamente la protección (pruebas de potencia con desafío) (Pospisil *et al.*, 1996; Vallat, 2004) o cuantificar el poder inmunogénico de las vacunas (Silva *et al.*, 2007b; Vallat, 2004). Si bien lo ideal es emplear en estas pruebas animales de la especie destino, que en este caso son los bovinos, los costos elevados que resultan de utilizar esta especie motivan el estudio de modelos con animales de laboratorio, que permiten obtener resultados extrapolables a un costo menor. En este sentido, Parreño *et al.* (2010a) han desarrollado un modelo empleando cobayos como especie alternativa que permitió concluir que las pruebas serológicas en cobayos podían emplearse para predecir la respuesta serológica y de protección en bovinos. Las conclusiones se basaron en el empleo correcto de pruebas estadísticas. En este trabajo se parte de esta experiencia previa para desarrollar un modelo estructural que contemple todas las variables intervinientes en una serie de ecuaciones de regresión múltiples distintas pero interrelacionadas, incluyendo variables que representan conceptos no observables en forma directa. A diferencia del trabajo mencionado que incluyó análisis univariados, el modelo propuesto incluye a todas las variables dependientes en un único modelo estructural.

Para probar la validez del modelo se debió recurrir a una simulación de la distribución multivariada de todas las variables observadas debido a que los datos disponibles constituían una muestra pequeña. Asimismo, en el diseño de los experimentos no se contemplaron muestras independientes de bovinos a ser empleadas para la evaluaciones de protección y serología, como se supuso en este modelo. En los experimentos originales los mismos bovinos que fueron evaluados serológicamente, fueron luego expuestos al virus en los ensayos de desafío y evaluados para determinar el grado de protección frente a la infección y la enfermedad por BoHV-1 causada en forma experimental, bajo condiciones controladas en boxes de bioseguridad. Es importante destacar que el tipo de vacunas utilizadas en nuestro país (vacunas a virus completo inactivados emulsionado en adyuvantes acuosos y oleosos) no otorgan inmunidad esterilizantes frente a la infección por este patógeno, por lo tanto todos los animales aún vacunados pueden infectarse luego del contacto con el virus pero, en presencia de anticuerpos vacunales inducidos por vacunas de calidad satisfactoria y muy satisfactoria, se espera una reducción significativa de la severidad de los síntomas clínicos y una reducción significativa en la curva de excreción viral en sus secreciones nasales.

En el modelo propuesto, los animales de ambas pruebas (serología y protección al desafío), aunque son vacunados con el mismo lote de vacuna a evaluar por calidad, no son los mismos, sino que constituyen muestras independientes. No obstante, para la simulación de las observaciones normales multivariadas se emplearon la matriz de covariancias y el vector de medias estimados a partir de la muestra. Como se partió de una distribución Normal multivariada, el método de estimación de parámetros seleccionado fue el de Máxima verosimilitud ya que se tuvieron en cuenta las ventajas que este método ofrece y que fueron enumeradas en el Capítulo 2.

La validez del modelo estructural final depende del establecimiento de la fiabilidad de los constructos subyacentes (Schreiber *et al.*, 2006). Se puede apreciar

mejor las potencialidades del modelo estructural cuando primero se prueban las variables indicadoras para cada variable latente en un modelo de medida que se empleará en el análisis del modelo estructural final. Por esta razón, en este trabajo, se siguió el análisis del modelo en dos etapas. En una primera etapa (Sección 1 del Capítulo 4) se analizó un modelo de medida en el que las variables latentes se correlacionaron entre sí, considerándolas variables exógenas, y en una segunda etapa, se consideró a la variable latente Protección a la enfermedad como endógena, causada por la variable Protección a la infección. Si bien, al ser inoculados con el virus, todos los animales se infectaron, se espera que aquellos animales que eliminan/excretan menos el virus, presentarán síntomas clínicos de menor severidad que los animales controles no vacunados y desafiados con la misma dosis viral.

En el análisis del modelo de medida tal como se propuso en el Capítulo 3, se observó un ajuste global pobre. El valor del estadístico χ^2 fue muy alto, algo superior a 20 veces los grados de libertad. Si bien existen ciertas consideraciones en el empleo de la prueba de hipótesis por χ^2 que fueron discutidas en el Capítulo 2, los índices de ajuste descriptivos, estuvieron muy por debajo del límite 0.90 considerado aceptable, indicando, también, un ajuste pobre. No obstante todos los coeficientes correspondientes a las variables indicadoras fueron significativos, con valores de estadísticos t superiores a 5, por lo que se pudo confirmar que todas las variables observadas estaban midiendo los constructos que se suponía debían medir.

Cuando se analizaron los índices de modificación de Lagrange, se observó, que podía obtenerse una mejora importante con el agregado de un término de covarianza distinto de cero entre los errores correspondientes a las mediciones de anticuerpos en bovinos y cobayos por seroneutralización. Esta modificación tiene un fundamento teórico debido a la naturaleza del ensayo de determinación de anticuerpos por seroneutralización. Este es un ensayo biológico que involucra células susceptibles y virus infecciosos en el que se mezcla una dosis fija de virus con diluciones seriadas del suero del animal vacunado. Esta mezcla se incuba por una hora, momento en el cual los anticuerpos presentes en el suero neutralizan al virus e impiden la infección de las células que se agregan en un paso posterior. El ensayo se revela detectando la presencia de efecto citopático viral en las células, (es decir, la presencia del virus residual no neutralizado). La lectura del ensayo se realiza por un operador calificado. Si bien el ensayo cuenta con controles de cada uno de sus componentes, posee alto grado de subjetividad que depende del estado de las células, el rango de dosis viral admisible para tomar el ensayo como válido y el criterio del observador. Finalmente, sueros con títulos altos de anticuerpos son menos susceptibles a las variaciones que los sueros con títulos bajos. De lo expuesto se desprende que los valores de seroneutralización son altamente dependientes del día de medición. Pequeños cambios en estas variables podrían reflejarse en lecturas diferentes, por lo que medidas hechas en un mismo día o mismas condiciones pueden resultar en errores correlacionados. En este aspecto, las mediciones por ELISA si bien no están libres de este problema, son menos sensibles a variaciones en este tipo de variables dado que es un ensayo que ha sido validado bajo normas ISO 17025 y presentó coeficientes de variación dentro del rango admitido por la OIE (Parreño *et al.*, 2010b).

La posibilidad de agregar un término de correlación de errores distinto de 0 para las variables de seroneutralización constituye una ventaja de la técnica de ecuaciones estructurales frente a otras técnicas multivariadas como el análisis factorial exploratorio. El modelo se modificó de modo que la matriz Θ_8 dejó de ser una matriz diagonal. El ajuste del modelo mejoró significativamente al quitar esta restricción. No obstante, las medidas de diagnóstico del modelo siguieron indicando un ajuste global pobre. A partir

de los índices de Lagrange de este nuevo modelo, se observó que se podría obtener una mejora en el ajuste quitando la restricción de covariancia igual a cero entre errores que involucraban el error de medición de la variable Tiempo hasta el día pico de severidad del cuadro clínico, indicadora de la Protección a la enfermedad. Este error aparecía en 6 de los 10 multiplicadores de Lagrange mayores, indicando una mejora en el ajuste sacando la restricción de no correlación (Cuadro 4.9 del Capítulo 4). Si bien la variable Tiempo hasta el día pico de severidad del cuadro clínico había mostrado ser indicadora de la Protección a la enfermedad con coeficiente significativamente distinto de 0, el porcentaje de variabilidad de esa variable que fue explicado por la variable latente fue muy bajo (16%) por lo que se consideró la eliminación de esta variable indicadora del modelo, teniendo en cuenta que quedaban 4 variables indicadoras más de la Protección a la enfermedad. De este modo, no se comprometía la identificación del modelo.

Con la segunda modificación del modelo, el ajuste mejoró significativamente, pero aún no fue satisfactorio. Dentro de los multiplicadores de Lagrange mayores indicando mejoras con el agregado de componentes, se presentaron covariancias de errores correspondientes a las variables indicadoras de la Protección a la infección, Pico de excreción y Tiempo al día pico de excreción. Como la variable Tiempo al día pico de excreción, si bien fue significativa, presentó baja fiabilidad (10%) se decidió eliminar esta variable del modelo. El ajuste del modelo mejoró con esta modificación.

Un problema no abordado hasta aquí fue la presencia de valores imposibles en la estimación del modelo. Se presentaron 2 coeficientes estandarizados mayores a 1 y, por lo tanto, 2 variancias de error menores a 0. El problema de las soluciones impropias de un modelo de ecuaciones estructurales estimado es conocido como “caso Heywood” y es frecuente en este tipo de modelos. Dada la imposibilidad de estos valores en la población, se hace necesario el análisis de la causa de este problema. Existen trabajos estadísticos dedicados a la comprensión de las causas probables de las soluciones impropias. Entre otros, Chen *et al.* (2001), Gagne y Hancock (2006), Gerbing y Anderson (1987), abordan el tema. Valores estandarizados de coeficientes superiores a 1 pueden ser válidos algunas veces (Joreskog, 1999), pero, también pueden indicar que se han impuesto restricciones no razonables al modelo o que existen otros problemas en el modelo. No existe una única causa del caso Heywood. Existe amplia evidencia de que el peligro real de ocurrencia de soluciones impropias es una muestra pequeña, pero hay otros factores que inciden. En este sentido, Boomsma y Hogland (2001) demostraron, a partir de simulaciones de Monte Carlo, que matrices de covariancia poblacionales con valores crecientes de variancias presentaban menor proporción de soluciones impropias, esto es, que indicadores con valores mayores de coeficientes en el modelo presentaban mayor proporción de soluciones impropias. Este fenómeno es conocido como “caso cercano a cero” (Van Driel, 1978), en la medida en que las variancias poblacionales se acercan a cero, la probabilidad de obtener valores negativos de esas variancias aumenta. Otro factor que incide en la probabilidad de obtener soluciones impropias es el valor del cociente entre el número de variables indicadoras y el número de variables latentes, con valores mayores de este cociente produciendo menor proporción de soluciones impropias. Estos mismos factores inciden en la convergencia del modelo y, pueden ser síntomas de subidentificación empírica (Boonsma y Hogland, 2001). En estos casos, debido a que hay poca cantidad de información (traducida en términos de observaciones independientes, la fiabilidad y validez de las medidas), es más probable que ocurran inconsistencias entre un modelo hipotético y los datos empíricos. En este trabajo, como no se detectaron otros problemas como la no convergencia, o errores estándares de los estimadores muy bajos, se decidió restringir el valor de las variancias del error negativas a un valor muy pequeño (0.0000001). Los coeficientes que superaban a 1

correspondieron a las variables de áreas bajo la curva de excreción y de síntomas clínicos, indicadoras de las variables latentes Protección a la infección y Protección a la enfermedad, respectivamente, que eran las variables indicadoras que se esperaba fueran las mejores indicadoras de esos constructos.

Hasta aquí, la respuesta al ajuste global pobre fue la re-especificación del modelo con pocos cambios que permitieron una mejora significativa en el ajuste. Sin embargo, el ajuste global del modelo no alcanzó a ser satisfactorio de acuerdo a los índices de ajuste calculados. La estrategia sugerida por Bollen (1989) de analizar el modelo por partes para identificar cuál es la fuente de falta de ajuste, arrojó cierta luz al problema. Por un lado, cada variable latente por separado mostró un buen ajuste global, siendo la variable Calidad, según el valor obtenido en el modelo animal de laboratorio, la que presentó el mejor ajuste.

Al considerar las variables latentes de a pares, se observó que el ajuste fue satisfactorio para todos los pares de variables excepto para la Protección a la Infección y Protección a la enfermedad. Una de las posibles causas de la especificación imperfecta del modelo en este caso, podría ser la restricción impuesta de no correlación de los errores de medición de una ecuación con los errores de otra en el caso de las variables de protección. El empleo del mismo animal en la evaluación de las variables indicadoras de una misma variable latente, pueden llevar a la necesidad de incluir términos de correlación de errores distintos de 0. La posibilidad de incluir errores correlacionados en el modelo constituye una de las ventajas de este tipo de modelos como se mencionó antes. Sin embargo, la mayor complejidad de este nuevo modelo teórico puede llevar a una subidentificación del modelo. Por otra parte, la re-especificación basada en los resultados del modelo inicial transforma el análisis de confirmatorio en exploratorio. Como consecuencia, los niveles de probabilidad de las pruebas estadísticas deben ser considerados aproximados (Bollen, 1989). **La fuente más rica de ideas para la re-especificación debe ser el conocimiento teórico del fenómeno y no los resultados empíricos.**

En este caso, si se re-especificara el modelo agregando correlaciones de errores de las variables indicadoras de protección, se debería probar el ajuste del nuevo modelo en una nueva muestra independiente para evitar que sean los datos los que guíen la modificación del modelo teórico.

Los resultados hasta aquí indicaron que el modelo tal como se planteó en un principio no permite reproducir exactamente la matriz de covarianza poblacional (Σ) a partir de la matriz de covarianza en función de los parámetros ($\Sigma(\theta)$). Como se mencionó, la principal fuente de falta de ajuste se encuentra en la componente que incluye a las dos variables de Protección. **Es interesante notar que, cuando se consideró el modelo sólo con las variables Protección a la infección y Calidad por laboratorio, el ajuste fue bueno.** Asimismo, se observó un buen ajuste cuando se consideraron las variables latentes individualmente. En este sentido, uno de los principales intereses en la evaluación final de la calidad de la vacuna es la comprobación de la validez y confiabilidad de las variables indicadoras de las distintas dimensiones de la calidad. Los valores de los coeficientes estandarizados estimados fueron altos para todas las variables indicadoras de la Protección a la infección y la Calidad por laboratorio. Los valores de las variables indicadoras de la Protección a la enfermedad, Pico de excreción y Puntaje a los 14 días, fueron los más bajos. Sin embargo, todos los coeficientes fueron significativos, indicando que todas las variables indicadoras estaban midiendo el constructo correspondiente. **Los valores de fiabilidades más altas de las variables indicadoras correspondieron a las variables latentes Calidad por laboratorio, y dentro de estas se destacaron las variables de**

anticuerpos medidas por ELISA, ya sea en bovinos (0.92) o en cobayos (0.84). Las medidas de fiabilidad compuesta fueron altas para las 3 variables latentes, lo que indicó una alta consistencia de las variables indicadoras en medir esos constructos. Estos índices fueron de alrededor de 0.90 para la Protección a la infección y Calidad por laboratorio, y fue algo más bajo en la Protección a la enfermedad (0.82). Los índices de Variancia extraída, que miden cuánto de la variabilidad de las variables indicadoras es explicada por el constructo, en relación al error de medida, fueron altos para las dos primeras variables latentes (0.75 y 0.78 para la Protección a la infección y Calidad por laboratorio, respectivamente), y fue moderadamente bajo, pero aceptable, para la Protección a la enfermedad (0.55). **Las variables Área bajo la curva de excreción y Área bajo la curva de severidad demostraron ser las mejores indicadoras de la Protección a la infección y Protección a la enfermedad, respectivamente, confirmando lo observado por Parreño *et al.* (2010a).**

Como se ha mencionado, el modelo teórico consta de 2 componentes: el modelo de medida, que describe las relaciones entre las variables latentes y sus indicadoras, y un modelo estructural que describe las relaciones entre las variables latentes. En este trabajo, se postuló que **las variables protección a la Infección y Calidad por laboratorio están correlacionadas y que la Protección a la enfermedad se da en la medida en que se logra reducir la severidad de la infección. Todos los animales vacunados pueden infectarse, pero al aumentar la calidad de la vacuna y el nivel de anticuerpos en los animales vacunados se registran infecciones con severidad clínica baja hasta llegar a ser subclínica (ausencia de enfermedad aparente).** Todos los índices de ajuste global analizados hasta aquí reflejan el ajuste del modelo completo, el modelo de medida y el modelo estructural combinados. Sin embargo, la porción estructural del modelo completo está formado sólo por 3 variables latentes y, por otro lado, hay 11 variables indicadoras. Esto significa que la porción de medida del modelo contiene muchos más parámetros a estimar que la porción estructural del modelo. En consecuencia, los índices de ajuste global están mucho más influenciados por el ajuste de la componente de medida que de la estructural. De hecho, Mulaik *et al.* (1989) se refieren a situaciones en las que el modelo de medida provee un buen ajuste y, sin embargo, existen malas especificaciones en la porción estructural del modelo. Por esta razón, se calculó el índice de ajuste normado (RNFI), que refleja sólo el ajuste de esta parte del modelo. En este trabajo se obtuvo un RNFI superior a 0.91, indicando un buen ajuste y, por lo tanto, que la estructura de interrelación de las variables latentes fue correcta.

El coeficiente de correlación entre las variables latentes Protección a la infección y Calidad por laboratorio fue alto (0.83), indicando que ambos constructos están muy relacionados entre sí. **Este resultado es relevante ya que, uno de los principales objetivos en el planteo del modelo era verificar que los valores serológicos obtenidos en el laboratorio se relacionaban con la protección del bovino. Es interesante notar, que dentro de los valores de laboratorio se encuentran los niveles de anticuerpos obtenidos en la otra especie, los cobayos, que mostraron altas fiabilidades, en especial los valores de la variable de anticuerpos medidos por la técnica de ELISA.**

El coeficiente que mide el efecto directo de la Protección a la infección en la Protección a la enfermedad, si bien fue significativo, fue bajo (0.53), dando una fiabilidad de 0.28. Si bien este resultado podría sorprender en un primer momento, un análisis más cuidadoso estaría indicando que los signos y síntomas evaluados en la Protección a la enfermedad pueden ser inespecíficos para el virus de la Rinotraqueitis. Por otra parte, como se ha mencionado en el Capítulo 1, este virus puede constituirse en

puerta de entrada para infecciones secundarias. Por otra parte, dado que la infección puede mantenerse en forma subclínica en el animal, esto es sin la manifestación de síntomas, existe la posibilidad de que animales con distintos niveles de infección no presenten signos visibles de enfermedad (esto es lo máximo que se puede lograr en este tipo de vacunas a virus inactivado). Este hecho estaría indicando la necesidad de considerar a la Protección a la infección en la evaluación de las vacunas contra BoHV-1, o a las variables serológicas evaluadas en el laboratorio a partir de animales no inoculados con el virus. **Asimismo, las evaluaciones en cobayos se presentan como una alternativa posible para evaluar la calidad de la vacuna.**

Capítulo 6: CONCLUSIONES

El modelo de ecuaciones estructurales demostró ser una técnica útil para la validación de un modelo empleando mediciones serológicas en animales de laboratorio como alternativa a las clásicas pruebas de desafío en bovinos, para la evaluación de la efectividad de las vacunas utilizadas en el control y prevención de la enfermedad causada por un herpesvirus.

En los ensayos de desafío, en los que los bovinos son inoculados con el virus luego de la vacunación, resulta difícil evaluar el grado de protección logrado por la vacuna. Se requieren normalmente varias variables para medir este concepto. Asimismo, es difícil de asumir que esas variables están medidas sin error. Se sabe, tanto desde la perspectiva práctica como de la teórica, que no se puede medir perfectamente un concepto y que siempre hay algún grado de error de medida. El modelo de ecuaciones estructurales abordó estos problemas representando los conceptos no observados en variables latentes y teniendo en cuenta el error de medida en el proceso de estimación. En este trabajo se pudo comprobar que las variables que representaban los conceptos de interés, Protección a la infección, Protección a la enfermedad y Calidad determinada por el modelo animal en el laboratorio (medición alternativa de la efectividad de la vacuna) explicaron adecuadamente la variabilidad de las variables observables. Se destacaron las variables Áreas bajo las curvas de excreción y severidad para representar las variables Protección a la infección y a la enfermedad, respectivamente, confirmando lo observado por Parreño *et al.* (2010a), y las variables serológicas medidas por ELISA en bovinos y cobayos representando la Calidad por laboratorio. Por otra parte, el modelo permitió la incorporación de correlación de errores de las variables serológicas, que hubiese sido imposible con otras técnicas multivariadas.

La técnica de análisis presentada en este trabajo, permitió explorar por partes las relaciones entre variables para detectar la fuente de falta de ajuste. En esta etapa de análisis se pudo comprobar el buen ajuste de las relaciones de las variables latentes Protección a la infección y Calidad por laboratorio, obteniéndose una alta correlación entre ambos conceptos. Teniendo en cuenta la alta fiabilidad de las variables serológicas medidas por ELISA tanto en bovinos como en cobayos, se puede inferir que la medición de estas variables de laboratorio puede ser útil en la predicción de la protección de la vacuna. El efecto de la Protección a la infección sobre la Protección a la enfermedad fue significativo pero débil. Este resultado podría deberse a que, en ocasiones, el animal puede estar inmunizado en diverso grado y presentar una baja protección a la infección, pero, dadas las características del virus y el desarrollo de respuesta inmune inducida por las vacunas, el animal vacunado e infectado luego de la exposición al agente puede no presentar síntomas (alta protección a la enfermedad) y estar excretando virus en forma subclínica. Sin embargo, si el animal excreta virus y lo hace por más tiempo (baja protección a la infección), va a ser un vehículo de contagio para mayor cantidad de animales en el rodeo. Por esta razón, se busca una vacuna que, fundamentalmente, proteja contra la infección, es decir proteja frente al cuadro clínico y reduzca al mínimo la excreción viral, si bien en este caso la vacuna nunca llegará a ser esterilizante (protección total a la infección, como sí se logra para otros virus con fiebre aftosa).

Otra de las virtudes del análisis de ecuaciones estructurales es la posibilidad de evaluar, por separado, el aporte del modelo de medida y del estructural en el ajuste

global del modelo. En este trabajo se comprobó que la parte estructural fue adecuada, validando la teoría inicial de interrelaciones entre las variables latentes.

Es importante destacar que la técnica de modelación por ecuaciones estructurales es una técnica confirmatoria y no exploratoria a diferencia de otras técnicas multivariadas clásicas. En este trabajo, se partió de un modelo que había sido validado por otras técnicas estadísticas y que fue verificado con la técnica propuesta, lográndose, con esta última, profundizar en las interrelaciones entre variables y considerar los errores de medida.

El modelo de ecuaciones estructurales podría extenderse a otros sistemas de evaluación de vacunas para otros virus. Por ejemplo, para rotavirus, principal agente causal de diarreas en terneros, se deberían considerar algunas otras variables debido a que la inmunidad del ternero es conferida por transferencia de anticuerpos pasivos de la madre a su cría por la ingestión de calostro y leche. El investigador debería contemplar estas diferencias para la formulación del modelo teórico.

Una de las desventajas del empleo del modelo de ecuaciones estructurales en este tipo de estudios podría ser el requerimiento del tamaño de muestra, especialmente en lo que se refiere a las pruebas de desafío. En este trabajo se emplearon simulaciones basadas en la información de experimentos realizados por el INTA. Sin embargo, si se considera la disminución de ensayos de desafío que involucraría la adopción de un método de laboratorio alternativo validado con cobayos para evaluar un lote de vacuna, el diseño de un ensayo que contemple un mayor número de animales inoculados con el virus sería de gran beneficio para la ganadería de nuestro país.

APÉNDICE

A.1. Programa para simular los valores serológicos de cobayos a partir de los valores bovinos

```
data a.corre;  
  
    set a.regres;  
  
    x8s= -0.22 + (1.11*x6) + (rannor(14575647)*0.173);  
  
    x7s= 0.14 + (0.82*x5) + (rannor(11287578)*0.447);  
  
run;  
  
proc print;  
  
    run;  
  
proc reg data= a.corre;  
  
    model x8s= x6;  
  
    model x7s= x5;  
  
run;
```

A2 Programa para generar la muestra Normal multivariada

```
%macro mvn(varcov=, means=, n=, myMVN=);  
  
/* arguments for the macro:  
  
1. covcov: data set for variance-covariance matrix  
  
2. means: data set for mean vector  
  
3. n: sample size  
  
4. myMVN: output data set name */  
  
proc iml;
```

```

use &varcov; /* read in data for variance-covariance matrix */

read all into sigma;

use &means; /* read in data for means */

read all into mu;

p = nrow(sigma); /* calculate number of variables */

n = &n;

l = t(half(sigma)); /* calculate cholesky root of cov matrix */

z = normal(j(p,&n,1234)); /* generate nvars*samplesize normals */

y = l*z; /* premultiply by cholesky root */

yall = t(repeat(mu,1,&n)+y); /* add in the means */

varnames = { x1 x2 x3 x4 x5 x6 x7 x8 y1 y2 y3 y4 y5 };

create &myMVN from yall (|colname = varnames|);

append from yall;

quit;

%mend mvn;

data means1;

input x @@;

cards;

-6.17 -5.47 -5.32 -29.08 2.08 1.41 1.90 1.44 -2.79 -5.44 -9.12 -11.62 -19.31

;

run;

data varcov1;

input x1-x8 y1-y5;

cards;

1.1294118 0.1475936 0.9310160 9.2411943 0.8087504 0.5163323 0.6633886
0.6112816 0.1755793 -0.6775401 -0.3267380 1.4217469 0.5896613

0.1475936 0.7415330 0.6007130 3.1008913 0.1712656 0.1883487 0.2163280
0.2168271 0.0998217 1.4527629 -0.0267380 -0.5418895 1.1532977

```

0.9310160 0.6007130 2.9527629 15.6462567 0.8418360 0.6770649 0.8928164
0.7689929 0.3716578 1.6711230 2.5971480 3.2183601 8.1546346

9.2411943 3.1008913 15.6462567 128.6122906 9.2936435 6.0434087 8.7843993
6.8100517 2.8501783 6.9426916 16.9418895 25.2888592 55.8603387

0.8087504 0.1712656 0.8418360 9.2936435 1.0687544 0.5828799 0.9430537
0.6624358 0.1446168 -0.3511052 0.8246346 1.9337255 2.5079234

0.5163323 0.1883487 0.6770649 6.0434087 0.5828799 0.5012730 0.5440411
0.5517419 0.1318746 -0.0369743 0.4791144 1.3322534 1.7862539

0.6633886 0.2163280 0.8928164 8.7843993 0.9430537 0.5440411 0.9947258
0.6054640 0.1567201 -0.0643137 0.9977184 1.6419608 2.4137077

0.6112816 0.2168271 0.7689929 6.8100517 0.6624358 0.5517419 0.6054640
0.6186514 0.1355169 -0.0598307 0.4042068 1.3976916 1.6533913

0.1755793 0.0998217 0.3716578 2.8501783 0.1446168 0.1318746 0.1567201
0.1355169 0.4714795 0.7299465 0.6916221 1.4643494 3.5200535

-0.6775401 1.4527629 1.6711230 6.9426916 -0.3511052 -0.0369743 -0.0643137 -
0.0598307 0.7299465 8.3146168 4.5828877 -0.7959002 11.9656863

-0.3267380 -0.0267380 2.5971480 16.9418895 0.8246346 0.4791144 0.9977184
0.4042068 0.6916221 4.5828877 13.3190731 5.9857398 28.4625668

1.4217469 -0.5418895 3.2183601 25.2888592 1.9337255 1.3322534 1.6419608
1.3976916 1.4643494 -0.7959002 5.9857398 14.8493761 22.8943850

0.5896613 1.1532977 8.1546346 55.8603387 2.5079234 1.7862539 2.4137077
1.6533913 3.5200535 11.9656863 28.4625668 22.8943850 88.5456774

;

run;

%*mvn*(varcov=varcov1, means=means1, n=**200**, myMVN=base1)

data base1;

set base1;

array ar1(**4**) x5-x8;

array ar2(**9**) x1-x4 y1-y5;

do i=**1** to **4**;

if ar1(i)<**0.3** then ar1(i)=**0.3**;

```

end;
do j=1 to 9;
if ar2(j)>0 then ar2(j)=0;
end;
run;

```

A3. Programa para el análisis del modelo de ecuaciones estructurales para la validación de un método alternativo de evaluación de vacunas

```
/*Modelo de medida*/
```

```
proc calis DATA=base1 COVARIANCE CORR RESIDUAL MODIFICATION;
```

```
LINEQS
```

```
x1= Lx1F1 F1 + E1,
```

```
/*x2= Lx2F1 F1 + E2,*/
```

```
x3= Lx3F1 F1 + E3,
```

```
x4= Lx4F1 F1 + E4,
```

```
x5= Lx5F3 F3 + E5,
```

```
x6= Lx6F3 F3 + E6,
```

```
x7= Lx7F3 F3 + E7,
```

```
x8= Lx8F3 F3 + E8,
```

```
y1= Ly1F2 F2 + Ey1,
```

```
/*y2= Ly2F2 F2 + Ey2,*/
```

```
y3= Ly3F2 F2 + Ey3,
```

```
y4= Ly4F2 F2 + Ey4,
```

```
y5= Ly5F2 F2 + Ey5;
```

```
STD
```

F1=1,

F2=1,

F3=1,

E1 /*E2*/ E3 E5-E8= VARE1 /*VARE2*/ VARE3 VARE5-VARE8,

E4= 0.0000001,

Ey1 /*Ey2*/ Ey3 Ey4 /*Ey5*/= VAREy1 /*VAREy2*/ VAREy3 VAREy4
/*VAREy5*/,

Ey5= 0.0000001;

COV

E8 E6= CE8E9,

/*E10 E11= CE10E11,

E8 E10= CE8E10,

E9 E11= CE9E11,*/

F1 F2=CF1F2,

F1 F3=CF1F3,

F2 F3=CF2F3;

VAR x1 /*x2*/ x3-x8 y1 /*y2*/ y3 y4 y5;

RUN;

/*Considerando la estructura de las variables latentes*/

proc calis DATA=base1 COVARIANCE CORR RESIDUAL MODIFICATION;

LINEQS

x1= Lx1F1 F1 + E1,

/*x2= Lx2F1 F1 + E2,*/

x3= Lx3F1 F1 + E3,

x4= F1 + E4,

x5= F3 + E5,

$$x6 = Lx6F3 F3 + E6,$$

$$x7 = Lx7F3 F3 + E7,$$

$$x8 = Lx8F3 F3 + E8,$$

$$y1 = Ly1F2 F2 + Ey1,$$

$$/*y2 = Ly2F2 F2 + Ey2,*/$$

$$y3 = Ly3F2 F2 + Ey3,$$

$$y4 = Ly4F2 F2 + Ey4,$$

$$y5 = F2 + Ey5,$$

$$F2 = PF2F1 F1 + D1;$$

STD

$$F1 = VARF1,$$

$$F3 = VARF3,$$

$$E1 E3 E5-E8 = VARE1 VARE3 VARE5-VARE8,$$

$$E4 = \mathbf{0.0000001},$$

$$Ey1 /*Ey2*/ Ey3 Ey4 /*Ey5*/ = VAREy1 /*VAREy2*/ VAREy3 VAREy4 /*VAREy5*/,$$

$$Ey5 = \mathbf{0.0000001},$$

$$D1 = VARD1;$$

COV

$$E8 E6 = CE8E9,$$

$$/*E10 E11 = CE10E11,$$

$$E8 E10 = CE8E10,$$

$$E9 E11 = CE9E11,*/$$

$$F1 F3 = CF1F3;$$

$$VAR x1 /*x2*/ x3-x8 y1 /*-y2*/ y3 y4 y5;$$

RUN;

BIBLIOGRAFÍA

- Ackermann, M., Engels, M., (2006), Pro and contra IBR-eradication. *Veterinary microbiology* 113, 293-302.
- Aguilar Setién, J. A. 1987. El virus de la Rinotraqueítis Infecciosa Bovina (bovid herpesvirus I): Propiedades y vacunación. *Rev. Ciencia Veterinaria* 4: 161-207.
- Akaike, H. (1974) a new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19: 716-723.
- Akaike, H. (1987). Factor Analysis and AIC. *Psychometrika*, 52, 317–332.
- Alonzo, p.; Reolón, E.; Acuña, P.; Leaniz, R.; Puentes, R.; Lavarello, L.; Maisonnave, J. (2008) Evaluación de un modelo en cobayo (*Cavia porcellus*) para estudiar la inmunogenicidad de vacunas comerciales inactivadas contra Herpesvirus bovino. *Veterinaria* 44 (172): 9-15.
- Anderson, J. C.; Gerbing, D. W. (1984) The effects of sampling error on convergence, improper solutions and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, 49:155-173.
- Anderson, J. C.; Gerbing, D. W. (1988) Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103: 411-423.
- Armstrong, P. R.; Gaston, K. J.; Hanley, N. D.; Ruffell, R. J. (2009) Contrasting approaches to statistical regression in ecology and economics. *Journal of Applied Ecology* 46: 265-268.
- Austin, M. (2007) Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling* 20(26):1-19.
- Baba, K.; Shiraki, K.; Kanesaki, T.; Yamanishi, K.; Ogra, P.L.; Yabnuchi, H.; Takahashi, M. (1987) Specificity of skin test with varicella-zoster virus antigen in varicella-zoster and herpes simplex virus infections. *J. Clin. Microbiol.* 53:31-42.
- Belsley, D. A.; Kuh, E.; Welsh, R. E. (1980) Regression Diagnostics Identifying Influential data and Sources of Collinearity. Wiley, New York.
- Bentler, P. M.; Bonett, D. G. (1980) Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin*, 88: 588-600.
- Bentler, P.M. (1989), *EQS, Structural Equations, Program Manual*, Program Version 3.0, Los Angeles: BMDP Statistical Software, Inc.

- Bitsch, V. 1978. The P37/24 modification of the infectious bovine rhinotracheitis virus serum neutralization test. *Acta Vet. Scand.*, **19**, 497–505.
- Bollen, K. A. (1986). Sample size and Bentler and Bonett's nonnormed fit index. *Psychometrika*, 51:375-377.
- Bollen, K. A. (1987), 'Outliers and improper solutions: A confirmatory factor analysis example', *Sociological Methods and Research* 15, 375–384.
- Bollen, K. A. (1988). A new incremental fit index for general structural equation models. A paper presented at 1988 Southern Sociological Society Meetings. Nashville, Tennessee.
- Bollen, K. A. (1989). Structural Equations with Latent Variables. Wiley & Sons, Inc.
- Bollen, K. A. (1996a). An Alternative Two Stage Least Square (2SLS) Estimator for Latent Variable Equations. *Psychometrika* 61:109-21.
- Bollen, K. A. (1996b). A Limited-Information Estimator for LISREL. Models with and without Heteroscedastic Errors. In *Advance Structural equations Modeling Techniques*. Marcoulides and R. Schumacker Hillsdale, NJ: Lawrence Erlbaum.
- Bollen, K. A.; Bauer, (2004). Automating the Selection of Model-Implied Instrumental Variables. *Sociological Methods & Research*, 92(4):425-452.
- Bollen, k. A.; Liang, J. (1988) some properties of Hoelter's CN. *Sociological Research and Methods*, 16: 492-503.
- Boomsma, A. (1982) The Robustness of LISREL against small sample sizes in factor analysis models. In K. G. Jöreskog y H. Wold, eds. *Systems under Indirect Observation*, Part I. Amsterdam: North-Holland, pp. 149-173.
- Boomsma, A. (1983) *On the Robustness of LISREL (Maximun Likelihood Estimation) against Small Sample Size and Nonnormality*. Amsterdam: Sociometrics Research Foundation.
- Boomsma, A. ;Hoogland, J. J. (2001), 'The robustness of LISREL modeling revisited', *Structural Equation Modeling: Present and Future*.
- Bozdogan, H. (1987), "Model Selection and Akaike's Information Criterion (AIC): The General Theory and its Analytical Extensions," *Psychometrika*, 52, 345–370.
- Browne, M.W. (1974). Generalized least –squares estimators in the analysis of covariance structures. *South African Statistical Journal*, ed. D.M. Hawkins, New York: Cambridge University Press8:1-24.
- Browne, M.W. (1982). Covariance Structures. *Topics in Multivariate Analyses*, ed. D.M. Hawkins, New York: Cambridge University Press.

- Browne, M. W. (1984). Asymptotic distribution free methods in analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37: 62-83.
- Browne, M. W.; Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37, 62-83.
- Browne, M. W; Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.
- Campero, C. M.; Moore, D. P.; Odeón, A. C.; Cipolla, A. M.; Odriozola, E. (2003). Aetiology of Bovine abortion in Argentina. *Vet. Res. Commun*, 27(5): 359-69.
- Carmines, E. G.; McIver, J. P. (1981) Analyzing models with unobserved variables : analysis of covariance structures. En G. W. Bohrnstedt y E. F. Borgatta, eds., *Social Measurement: Current Issues*. Beverly Hills, CA: Sage, pp. 65-115.
- CFR (1985). Bovine Rinotracheitis Vaccine, killed virus, Administration, N.A.a.R., ed. US Government printing office, pp. 670-671. Amended at 55 CFR 35562, Aug. 31, 1990 and 56 CFR 66786, Dec. 26, 1991.
- CFR. (2008). Code of Federal Regulations. USDA:113.216.
- Chen, F., Bollen, K. A., Paxton, P., Curran, P. & Kirby, J. (2001), 'Improper solutions in structural equation models: Causes, consequences, and strategies', *Sociological Methods and Research* 25, 223–251.
- Coenders, G.; Batista-Foguet, J. M.; Satorra, A. (1995). Scale Dependence of the True Score MTMM Model. In Saris, W. E. & Münnich, Á. (Eds.) *The Multitrait-Multimethod Approach to Evaluate Measurement Instruments*. Eötvös University Press, Budapest: 71-87.
- Correa, M. T.; Erb, H.; Scarlett, J. (1993) Path analysis for seven postpartum Disorders of Holstein cows. *J Dairy Sci* 76: 1305 – 1312.
- Cox, J.C.; Coulter, A. R. 1997. Adjuvants. A classification and review of their modes of action. *Vaccine*, 15(3): 248-256.
- Crow, J. F. (1988). Sewall Wright (1889-1988). *Genetics* 119(1):1-4.
- Davis, W. R. (1993) The FC1 rule of identification for confirmatory factor analysis: a general sufficient condition. *Sociological Methods & Research*, 21(4): 403-437.
- Del Medico Zajac MP, Puntel M, Zamorano PI, Sadir AM, Romera SA. BHV-1 vaccine induces cross-protection against BHV-5 disease in cattle. *Res Vet Sci* 2006;81(3):327–34.

- Del Médico Zajac MP, Romera SA, Ladelfa MF, Kotsias F, Thiry J, Ziant D, Meurens F, Keil GM, Thiry E, Muylkens B. (2009). Characterization of interspecific recombinants generated from closely related bovine herpesviruses 1 and 5 through multiple PCR sequencing assays. *J Virol Methods* 161(1):75-83
- De los Campos, G.; Gianola, D.; Heringstad, B. (2006a) A structural equation model for describing relationships between somatic cell score and milk yield in first-lactation dairy cows. *J. of Dairy Sc.* 89(11): 4445-4455.
- De los Campos G, Gianola D, Boettcher P, Moroni P. (2006b). A structural equation model for describing relationships between somatic cell score and milk yield in dairy goats. *J Anim Sci* 84:2934-2941.
- EMA, T.E.A.f.t.E.o.M. (1998). CVMP/IWP/P038/97-Position paper on batch potency testing of immunological veterinary medical products, Unit, V.M.E., ed. (EMA).
- EMA, T.E.A.f.t.E.o.M. (1999). EMA/CVMP/140/97-POSITION PAPER ON COMPLIANCE OF VETERINARY VACCINES WITH VETERINARY VACCINE MONOGRAPHS OF THE EUROPEAN PHARMACOPOEIA, Unit, V.M.E., ed. (Committee for veterinary medical products).
- Fornell, C.; Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing Research*, 18: 39-50.
- Gagne, P. E.; Hancock, G. R. (2006). Measurement model quality, sample size, and solution propriety in confirmatory factor models. *Multivariate Behavioral Research*, 41, 65-83.
- Gerbing, D. W.; Anderson, J. C. (1987). Improper solutions in the analysis of covariance structures: Their interpretability and a comparison of alternate respecifications. *Psychometrika*, 52, 99-111.
- Gianola D, Sorensen D: 2004. Quantitative genetic models for describing simultaneous and recursive relationships between phenotypes. *Genetics*, 167:1407-1424.
- Gibbs, E. P.J., Rweyemamu, M. M. (1997). Bovine herpesvirus. Part I. Bovine herpesvirus. *Vet. Bull.* 47:317-343.
- Gilliam S.E.; Thackray A.M.; Brown G.A.; Field H.J. (1993). The pathogenesis of wild type and drug resistant mutant strains of bovine herpesvirus-1 (BHV-1) in the natural host. *Arch Virol* 128:43-54.
- Goddard, R. D.; Hopkins, I. G.; Thornton, D. H. (1986) The development of a potency test for *Leptospira hardjo* vaccines: a comparison of protection in calves and serology in guinea-pigs. *J. Biol. Stand.* 14:337-344.

- Gómez, E.; Duato, L.; Dominguez Bernal, G.; Doménech Gómez, A.; Blanco Gutiérrez, M. del M.; Gibello Prieto, A.; Goyache Goñi, J.; Sánchez Vizcaíno Rodríguez, J. M.; Suárez Rodríguez, M.; Cutuli de Simón, M. T. (2011) Compendio de guiones para clases teóricas en la asignatura Inmunología en Veterinaria III: técnicas inmunológicas. *Reduca. Serie Veterinaria* 3 (15) 94-121.
- Gregoriadis G. (1990). Immunological adjuvants: a role for liposomes. *Immun. Today*, 11: 89-97.
- Haavelmo, T. (1944) The Probability Approach in Econometrics. *Econometrica* 12, iii-vi+1-115.
- Hage J.J., Schukken Y.H., Dijkstra T., Barkema H.W., Van Valkengoed P.H. & Wentink G.H. (1998). Milk production and reproduction during a subclinical bovine herpesvirus 1 infection on a dairy farm. *Prev. Vet. Med.*, **34**, 97–106.
- Hägglund, G. (1982) Factor analysis by instrumental variable methods. *Psychometrika*, 47:209-222.
- Hair, J. F.; Anderson, R. E.; Tatham, R. L.; Black, W. C. (1999) Análisis multivariante. Pearson-Prentice Hall, Madrid. 5ta Ed.
- Hair, J. F.; Anderson, R. E.; Tatham, R. L.; Black, W. C. (2007) Análisis Multivariante. Pearson- Prentice Hall, 5ta Ed. Madrid, España, pp. 611-690.
- Hatcher, L. (1994) A Step-by-Step Approach to Using SAS for Factor Analysis and Structural Equation Modeling. SAS Institute Inc., Cary, NC, USA
- Heringstad, B; Wu, X. L.; Gianola, D. (2008) Inferring relationships between health and fertility in norwegian red cows using recursive models. *J. Dairy Sci.* 92 :1778–1784.
- Hilgers LATH, Platenburg PLI, Luitjens A, Groenveld B, Dazelle T, Ferrari Laloux M, Weststrate MW. (1994). A novel non-mineral oil-based adjuvant. I. Efficacy of a synthetic sulfolipo-polysaccharide in a squalane-in-water emulsion in laboratory animals. *Vaccine*, 12:653-60.
- Hilgers LATH, Lejeune G, Nicolas I, Fochesato M, Boon B. (1999). Sulpholipo-cyclodextrin in squalane-in-water as a novel and safe vaccine adjuvant. *Vaccine* 17:219-28.
- Hoelter, J. W. (1983) The analysis of covariance structures: Goodness-of-fit indices. *Sociological Methods and research*, 11: 325-344.
- Institutional Animal Care and Use Committee Guidebook. (2002). ARENA/OLAW IACUC, 2nd Ed.

- Iriondo, J.M., Albert, M.J., Escudero, A., (2003). Structural equation modeling: an alternative for assessing causal relationships in threatened plant populations. *Biol. Conserv.* 113, 367–377.
- Jaganathan, K.S.; Rao, Y.U.; Singh, P.; Prabakaran, D.; Gupta, S.; Jain, A.S.; Vyas, P. (2005). Development of a single dose tetanus toxoid formulation based on polymeric microspheres: a comparative study of poly(D,L-lactic-co-glycolic acid) versus chitosan microspheres. *Int. J. Pharm.* 294, 23-32.
- James, L.R.; Mulaik, S.A.; Brett, J.M. (1982), *Causal Analysis: Assumptions, Models, and Data*, Beverly Hills: Sage Publications, Inc.
- Jöreskog, K. G. (1973). A general method for estimating a linear structural equation system. En A. S. Goldberger y O. D. Duncan, eds., *Structural Equation Models in the Social Sciences*. New York: Academic Press, pp. 85-112.
- Jöreskog, K. G. (1977). Structural equation models in the social sciences: specification, estimation and testing. In P. R. Krishnaiah, ed., *Application of Statistics*. Ámsterdam: North-Holland, pp. 265-287.
- Jöreskog, K. G.; Sörbom, D. (1986) LISREL VI: Analysis of Linear Structural Relationships by Maximum Likelihood and Least Square Methods. Mooresville, IN: Scientific Software, Inc.
- Jöreskog, K. G.; Sörbom, D. (1988) LISREL 7: *A Guide to the Program and Applications*. Chicago: SPSS, Inc.
- Joreskog, K. G., Sorbom, D. (1989). LISREL 7: A guide to the program and applications, 2nd Ed. Chicago: SPSS Inc.
- Joreskog, K.G. (1999). How large can a standardized coefficient be.. Informe no publicado. SSI Central, Inc.
<http://www.ssicentral.com/lisrel/techdocs/HowLargeCanaStandardizedCoefficientbe.pdf>
- Kahrs RF. (1977). Infectious Bovine Rhinotracheitis: A review and update. *JAVMA* 1977;171(10):1055-4.
- Kaplan, D. (2000). *Structural equation modeling: Foundation and extensions*. Thousand Oaks, CA: Sage Publications.
- Keesling, J. W. (1972) Maximum Likelihood Approaches to Causal Analysis. Ph. D. dissertation. Department of Education: University of Chicago.
- Kendrick J.W.; McEntree K. (1987). The effect of artificial insemination with semen contaminated with IBR–IPV virus. *Cornell Vet.*, **57**, 3–11.

- Kramps J.A.; Quak S.; Weerdmeester, K.; Van Oirschot J.T. (1993). Comparative study on sixteen enzymelinked immunosorbent assays for the detection of antibodies to bovine herpesvirus 1 in cattle. *Vet. Microbiol.*, **35**: 11–21.
- Kramps, J.A., Banks, M., Beer, M., Kerkhofs, P., Perrin, M., Wellenberg, G.J., Oirschot, J.T., (2004), Evaluation of tests for antibodies against bovine herpesvirus 1 performed in national reference laboratories in Europe. *Veterinary microbiology* 102, 169-181.
- Leenaars PPAM, Hendriksen Coenraad FM, de Leeuw Wim A, Florina Carat, Philippe Delahaut, Rene Fischer, et al. The production of polyclonal antibodies in laboratory animals. *ATLA* 1999;27:79–102.
- Lingling, H. 2006. Generating Multivariate Normal Data by Using PROC IML
Disponibile en: http://analytics.ncsu.edu/sesug/2006/CC15_06.PDF.
- López de Maturana E, Legarra A, Varona L, Ugarte E: 2007. Analysis of fertility and dystocia in Holsteins using recursive models to handle censored and categorical data. *J Dairy Sci*, 90:2012-2024.
- López de Maturana E, de los Campos, G.; Wu, X. L.; Gianola, D.; Weigel, K. A; Rosa, G. J. (2010). Modelling relationships between calving traits: a comparison between standard and recursive mixed models. *Genet Sel. Evol.* 42(1):1
- López, M. V.; Marangunich, L.; Bartoloni, N.; Parreño, V.; Rodriguez, D.; Vena, M. M.; Izuel, M. A comparison of classification tree and linear regression analysis for the assessment of vaccine quality. Presentación oral en el 56th Session of the International Statistical Institute (ISI) del 22 al 29 de Agosto de 2007 en Lisboa, Portugal. Book of Abstracts, pp. 274.
- Marquez, A. L.; Real, R.; Vargas, J. M. (2004) dependence of broad-scale geographical variation in fleshy-fruited plant species richness on disperser bird species richness. *Global Ecol. Biogeogr.* 13: 295-304.
- McCune, B.; Grace, J.B.; (2002) Analysis of Ecological Communities. MJM Software Design, Oregon, USA.
- McDonald, R. P. (1967a) Numerical methods for polynomial models in non-linear factor analysis. *Psychometrika*, 32:77-112.
- McDonald, R. P. (1967b) Nonlinear factor analysis. *Psychometric Monograph No. 15*.
- McQuitty, S. (2004) Statistical power and structural equation models in business research. *Journal of Business Research*, 57(2):175-183.
- McVey, D.S., Galvin, J.E., Olson, S.C., (2003), A review of the effectiveness of vaccine potency control testing. *Int J Parasitol* 33, 507-516.

- Meeusen, E. N. T. ; Walker, J.; Peters, A.; Pastoret, P. P.; Jungersen, G. (2007) Current Status of Veterinary Vaccines. *Clin. Microbiol. Rev.* 20(3): 489-510.
- Miller JM, Whetstone CA, Bello LJ, Lawrence WC. (1991). Determination of ability of a thymidine kinase-negative deletion mutant of bovine herpesvirus-1 to cause abortion in cattle. *Am J Vet Res* 52(7):1038-43.
- Mitchell D. (1974). An outbreak of abortion in a dairy herd following inoculation with an intramuscular infectious bovine rhinotracheitis vaccine. *Can Vet J* 1974;15:148.
- Mitchell, R.J., (1992). Testing evolutionary and ecological hypotheses using path analysis and structural equation modelling. *Funct. Ecol.* 6, 123–129.
- Mitchell, R. J. (1994) Effects of floral traits, pollinator visitation, and plant size on *Ipomopsis* aggregate fruit production. *Am. Nat.* 143: 670-889.
- Moore, D.P., Campero, C.M., Odeon, A.C., Bardon, J.C., Silva-Paulo, P., Paolicchi, F.A., Cipolla, A.L., (2003), Humoral immune response to infectious agents in aborted bovine fetuses in Argentina. *Revista Argentina de microbiologia* 35(3), 143-148.
- Morein, B., K. Lövgren, S. Höglund y B. Sundquist. (1987). The ISCOM: an immunostimulating complex. *Immun. Today*, 8: 333-338.
- Mulaik, S. A.; James, L. R.; Alstine, J. V.; Bennett, N.; Lind, S.; Stilwell, D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, 105: 430-445.
- Murray, R. ; Cohen, P.; Hardegree, M.C. (1972). Mineral oil adjuvants: biological and chemical studies. *Ann Allergy* 30:146.
- Nettleton PF, Sharp JM, Herring JA. (1984). Infectious bovine rhinotracheitis virus excretion after vaccination, challenge and immuno-suppression. In: Wittman G, Gaskill RM, Rhiza HJ, editors. Latent Herpes Virus Infections in Veterinary Medicine. Boston: *Martinus Nijhoff*, p. 191-209.
- Odeón, A. C. S.; EJA; Paloma, E. J.; Leunda, M. R.; Fernández Sainz, I. J.; Pérez, S. E.; Kaiser, G. G.; Draghi, M. G.; Cetrá, B. M.; Cano, A. (2001). Seroprevalencia de la Diarrea Viral Bovina, Herpesvirus Bovino y Virus Sincicial respiratorio en Argentina. *Rev. Vet. Med.* 82(4): 216-20.
- Odeón, A. C. 2002. Vacunas para el control de las infecciones virales de los bovinos. Disponible en http://www.produccion-animal.com.ar/sanidad_intoxicaciones_metabolicos/infecciosas/bovinos_en_general/60-vacunas_para_infecciones_virales.pdf.
- O'Hagan, D.T. ; Singh, V. (2003) Microparticles as vaccine adjuvants and delivery systems. *Expert Rev. Vaccines* 2(2): 269-283.

- OIE, W.O.f.A.H., (2008a), PRINCIPLES OF VALIDATION OF DIAGNOSIS ASSAYS FOR INFECTIOUS DISEASES, In: OIE (Ed.) Manual of Diagnostic Tests and Vaccines for Terrestrial Animals. OIE, Paris, France, pp. 34-45.
- OIE. (2008b). Infectious Bovine Rhinotracheitis/Infectious Pustular Vulvovaginitis. In: Manual of diagnostic tests and vaccines for terrestrial animals. Paris, France: OIE; p. 752–67.
- OIE. (2010). Infectious Bovine Rhinotracheitis/ Infectious Pustular Vulvovaginitis. En: Manual of diagnostic tests and vaccines for terrestrial animals. Paris, France.
- Parreño, V; López; M. V.; Rodriguez, D.; Vena, M. M.; Izuel, M.; Filippi, J.; Romera, A.; Faverin, C.; Bellinzoni, R.; Fernandez, F.; Marangunich, L. (2010a) Development and Statistical Validation of a Guinea Pig model as a method to evaluate the Potency against Bovine Rhinotracheitis Virus (IBR) of Combined Veterinary Vaccines by ELISA and Virus Neutralization. *Vaccine* 28(13): 2539-2549.
- Parreño V, Romera SA, Makek L, Rodriguez D, Malacari D, Maidana S, Compaired D, Combessies G, Vena MM, Garaicoechea L, Wigdorovitz A, Marangunich L, Fernandez F. (2010b). Validation of an indirect ELISA to detect antibodies against BoHV-1 in bovine and guinea-pig serum samples using ISO/IEC 17025 standards. *J Virol Methods*, 169(1):143-53.
- Patel, J. R. (2005). Relative efficacy of inactivated bovine herpesvirus-1 (BHV-1) vaccines. *Vaccine* 23:4054-4061.
- Peek, L.J.; Middaugh, C.R.; Berkland, C. (2008). Nanotechnology in vaccine delivery. *Adv. Drug Delivery Rev.* 60: 915-928.
- Pospisil, Z.; Krejci, J.; Jinek, P.; Lany, P.; Zendulkova, D.; Cihal, P. (1996). Development of a disease control programme based on the use of an inactivated vaccine against infectious bovine rhinotracheitis. *Vet Microbiol* 53:199-206.
- Potter M.; McCardle, R.C. (1964). Histology of developing plasma cell neoplasia induced by mineral oil in Balb/c mice. *J Natl Cancer Inst* 33:497.
- Puertes Colorado, E. (2009). El desarrollo industrial del medicamento inmunológico veterinario. Discurso de ingreso como Académica Correspondiente. Academia de Farmacia de Galicia.
- Rajput Z.I.; Hu, S.; Xiao, C.; Arijó, A.G.. (2007). Adjuvant effects of saponins on animal immune responses. *J. Zheijang Univ. Sci. B.* 8(3) :153-161.
- Ramon G. (1924) Sur la toxine et l'anatoxine diphtériques. Pouvoir flocculant et propriétés immunisantes. *Ann. Inst. Pasteur*, 38, 1-10.

- Reilly, T. (1995) A necessary and sufficient condition for identification of confirmatory factor analysis models of complexity one. *Sociological Methods & Research*, 23(4): 421-441.
- Rhee YO, Kim JH, Namgoong S. (1987). Immunogenicity of ND-EDS 76-IBD combined oil adjuvant vaccine. Research reports rural development administration (Livestock and Veterinary). Republic of Korea. 29: pp. 209-212.
- Romera SA, Hilgers LA, Puntel M, Zamorano PI, Alcon VL, Dus Santos MJ, et al. Adjuvant effects of sulfolipo-cyclodextrin in a squalane-in-water and water-in-mineral oil emulsions for BHV-1 vaccines in cattle. *Vaccine* 2000;19(1):132–41.
- Rothenberg, T. J. (1971) Identification in parametric models. *Econometrica*, 39: 577-591.
- Schreiber, J. B.; Nora, A.; Stage, F.; Barlow, E. A.; King, J. (2006) reporting Structural equation Modeling and Confirmatory Factor Analysis Results: A review. *The Journal of Educational research*, 99(6): 323-337.
- Sclove, L.S. (1987). Application of Model-Selection Criteria to Some Problems in Multivariate Analysis. *Psychometrika*, 52, 333–343.
- Schleiss, M. R. (2008) Comparison of vaccine strategies against congenital CMV infection in the guinea pig model. *J. Clin. Virol.* 41:224-230.
- Schreiber, J. B.; stage, F. K.; King, J.; Nora, A.; Barlow, E. A.(2006). reporting Structural equation Modeling and Confirmatory Factor Analysis results: A Review. *The Journal of Educational Research*, 99(6): 323- 337.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, 6: 461–464.
- Shipley, B.; Lechowicz, M. J. (2000) The functional co-ordination of leaf morphology, nitrogen concentration, and gas exchange in 40 wetland species. *Ecoscience* 7: 183-194.
- SENASA. (2006) Resolución 216/2006. Disponible en:
<http://www.senasa.gov.ar/contenido.php?to=n&in=1029&io=4627>.
- SENASA (2006) Resolución 351/06. Disponible en:
<http://www.senasa.gov.ar/contenido.php?to=n&in=1029&io=9818>
- Silva, L. F. ; Diel, D. G. ; do Carmo Cliento, M.; Weiblen, R. (2007a) Guinea pigs as a model test of bovine herpesvirus type I and bovine viral diarrhea virus inactivated vaccines. *Ciencia Rural* 37:1060-1065.

- Silva, L.; Weiblen, R.; Flores, E. (2007b). Imunogenicidade de vacinas comerciais inativadas contra o herpesvírus bovino tipo 1. *Ciência Rural*, Santa Maria 37:1471-1474.
- Simon, H. (1952). On the Definition of Causal Relation. *The Journal of Philosophy* 49: 517-528.
- Simon, H. (1954). Spurious Correlation: A Causal Interpretation. *Journal of The American Statistical Association* 49: 467-479.
- Smee, D. F.; Leonhardt, J. A. (1994) Vaccination against bovine herpes mamillitis virus infections in guinea pigs. *Intervirology* 37: 20-24.
- Steiger, J.H. and Lind, J.C. (1980), “Statistically Based Tests for the Number of Common Factors,” paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.
- Studdert, M.J. (1994). Bovine herpesvirus. *En: Encyclopedia of Virology*, Webster R.G. & Granoff A., eds. Academic Press, London, UK, 155–158.
- Swaminathan, H.; Algina, J. (1978). Scale Freeness in Factor Analysis. *Psychometrika*, 43,581-583.
- Taffs, R.E., (2001), Potency tests of combination vaccines. *Clin Infect Dis* 33 Suppl 4, S362-366.
- Tanaka, J. S.; Huba, G. J. (1985) A fit index for covariance structure models Under arbitrary GLS estimation. *British Journal of Mathematical and Statistical Psychology*, 38: 197-201.
- Thiry J, Saegerman C, Chartier C, Mercier P, Keuser V, Thiry E (2008). Serological evidence of caprine herpesvirus 1 infection in Mediterranean France. *Vet. Microbiol.* 128: 261-268.
- Vallat, B. (2004). Infectious Bovine Rhinotracheitis/Infectious Pustular Vulvovaginitis, *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals O.I.E.* pp. 514-526.
- Van Driel, O. P. (1978), On various causes of improper solutions in maximum likelihood factor analysis. *Psychometrika* 43, 225–43.
- Van Drunen Littel-van den Hurf, S. (2006). Rationale and perspectives on the success of vaccination against bovine herpesvirus-1. *Veterinary Microbiology* 113: 275–282.
- Van Oirschot, J.T.; Straver P.J.; Van Lieshout J.A.H., Quak J.; Westenbrink F; Van Exsel A.C.A. (1993). A subclinical infection of bulls with bovine herpesvirus type 1 at an artificial insemination centre. *Vet. Rec.*, **132**: 32–35.

- Varona L, Sorensen D, Thompson R: Analysis of litter size and average litter weight in pigs using a recursive model. (2007). *Genetics*, 177:1791-1799.
- Wiley, D. E. (1973). The identification problem for structural equation models with unmeasured variables. In A. S. Goldberger and O. D. duncan, eds., *Structural Equation Models in the Social Sciences*. New York: Academic Press, pp. 69-83.
- Williams, L. J.; Holahan, P. J. (1994) Parsimony-Based Fit Indices for Multiple-Indicator Models. *Structural Equation Modeling* 1(2): 161-189.
- Wyler R, Engels M, Schwyzer M. (1989). Infectious bovine rhinotracheitis/vulvovaginitis BHV-1. In: Wytman G, Becker I, editors. Herpesvirus diseases of cattle, horses and pigs. Developments in veterinary virology. Boston: Kluwer, 1989. p. 1-72.
- World Organization for Animal Health, O., (2008), INFECTIOUS BOVINE RHINOTRACHEITIS/INFECTIOUS PUSTULAR VULVOVAGINITIS, In: OIE (Ed.) Manual of Diagnostic Tests and Vaccines for Terrestrial Animals. OIE, Paris, France, pp. 752-767.
- Wu, X. L.; Heringstad, B.; Chang, Y. M. ; de los Campos, G.; D. Gianola, D. (2007) Inferring Relationships Between Somatic Cell Score and Milk Yield Using Simultaneous and Recursive Models. *J. Dairy Sci.* 90:3508–3521.
- Wu, X. L.; Heringstad, B.; D. Gianola, D. (2008) Exploration of lagged relationships between mastitis and milk yield in dairy cows using a Bayesian structural equation Gaussian-threshold model. *Genetic Selection Evolution* 40(4):333-357.
- Yao Q., Qian P., Huang Q.; Cao Y.; Chen H. (2008) Comparison of immune responses to different foot-and-mouth disease genetically engineered vaccines in guinea pigs. *J. Virol. Methods* 147:143–150.
- Yates WDG. (1982). A review of Infectious Bovine Rhinotracheitis shipping fever pneumonia and viral-bacterial synergism in respiratory disease of cattle. *Can J Comp Med.* 46:225-63.
- Zhao, X.; Xi, J. (2011). The vaccines for Bovine Herpesvirus type 1: A review. *African Journal of Biotechnology*, 10(50): 10072-10075.